Title
Components of Coherence

Author
Sasaki, Kelsey

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

COMPONENTS OF COHERENCE

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Kelsey M. Sasaki

September 2021

The Dissertation of Kelsey M. Sasaki
is approved:

Professor Pranav Anand, co-chair

Assistant Professor Amanda Rysling, co-chair

Associate Professor Maziar Toosarvandani

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Components of Coherence

by

Kelsey M. Sasaki

This dissertation is about how discourse is comprehended in real-time, and how it is represented in memory. This project begins to bridge the gap between formal linguistic theories of discourse coherence and structure and psychological theories of discourse processing. I argue that combining the insights of these two strands can bring us to a deeper understanding of the core puzzles in discourse research, e.g., what determines whether or not a sequence of clauses or sentences seems like a coherent discourse? What factors help or hinder the on-line comprehension, the incremental interpretation, of discourse? How is discourse structured in memory, and to what degree does this structure resemble formal linguistic representations of discourse? I approach these questions via three sets of experiments. First, I investigate the role of temporal information in on-line discourse comprehension, give special scrutiny to the common assumption that progression is the default temporal relation in discourse processing. My results suggest that, if progression is any sort of default, it is not the sort that has a strong influence on on-line processing. The second set of experiments is situated in the context of Situation Model Theory, a prominent psychological theory of discourse. I investigate the roles of temporal and causal information in the construction

of discourse structure in memory. Based on my findings, I argue that the influence of temporal factors may not be as clear-cut as previous work has suggested. I also argue that causal factors have a relatively large role in this domain. Finally, I turn to Segmented Discourse Representation Theory, a prominent linguistic coherence theory, and its Right Frontier Constraint, which is argued to be the primary governor of pronominal anaphora in discourse. I challenge this view, arguing that the effects that are attributed to this constraint can in fact be accounted for by independent factors that influence the salience of entities in discourse.

For Nancy and Greg Sasaki

Acknowledgments

Words cannot fully describe the gratitude and appreciation I have for all the people who have supported me throughout my time in graduate school, but they will have to do. To all of you: thank you for letting me stand on your shoulders, lean against your shoulders, and exist alongside your shoulders. I hope that you in turn have found, and will continue to find, my own shoulders worthy of standing on, leaning against, and existing alongside.

The thought of being a linguist in some place other than where the members of my committee are is genuinely terrifying to me. They have all given me unfailing patience and support throughout my years in Santa Cruz, though I'm sure I tested them all plenty of times. I am inexpressibly grateful for all the time and care they've invested in me.

Pranav Anand has been my advisor right from the start, although I came to Santa Cruz expecting to become a syntactician. This has been an unbelievable stroke of good fortune for me. Pranav's seemingly inexhaustible energy, curiosity, and generosity have inspired and influenced me in myriad ways, and likely always will.

Amanda Rysling is, in many ways, the perfect mentor for someone like me. If I grow up to be remotely as good of a coach as she is, I will be happy, though she will still be pushing me to be better, which is, of course, what makes her such a great coach. I am often surprised that our paths didn't cross sooner, but I am so glad that they eventually did.

Maziar Toosarvandani taught me most everything I know about fieldwork, and I am immeasurably the better for it. He has worked tirelessly to create fieldwork opportunities for us grads, and I have benefitted enormously from his efforts, both within and beyond the academic world. I will always treasure our shared experiences, including, in no particular order, building Nido, going to Costco, and teaching Laxopan schoolkids.

Thank you to Matt Wagers, for providing the foundation of my psycholinguistics education, and for all that he has done to help me build on it through z/lab.

Thank you to Jim McCloskey, for all his kind encouragement and guidance on my first QP, which turned out not to be a syntax project, and for all the lovely interactions we've had since then.

Thank you to Ashley Hardisty, Logan Roberts, Maria Zimmer, and Gwyn Vandevere, for all the effort and care they put into making the department run so beautifully. Special thanks to Maria and Logan for looking after us grad students so well.

Thank you to Fe Silva Robles, for welcoming me into her community, both in Santa Cruz and Laxopa. Working alongside her these past years has been an enormous privilege.

Thank you to Shayne Sloggett, for all the Ibex wizardry, stats-knowing, and general psycholinguistics hotline support. It has been no less appreciated for being at one degree of separation, and I hope to make a start of repaying it in the coming years.

After a couple of false starts as an undergrad, I was very fortunate to land in NYU's

linguistics department, which turned out to be exactly the right place for me. Thank you to Stephanie Harves, for being so welcoming and supportive, and the caliber of teacher I can only hope to become someday. To John Singler, whose Pidgins and Creoles class has had a profound impact on both my academic and personal lives, and whose adventurous and generous spirit continues to inspire me. To Chris Barker, who taught me what entailment is, and who gently advised me to 'Do what you want, but want Santa Cruz' when I was choosing a graduate program.

The strength of the graduate community here was a significant part of why I did indeed want Santa Cruz, and I do not believe I would've made it all the way through grad school without it. First, thank you to my cohort-mates, Tom Roberts and Jake Vincent. I'm so glad to have been on this journey with you both. Tom, I feel astoundingly lucky to have you as a friend, and the idea of moving far away is made much less intimidating by the knowledge that you will be moving not too far away from me.

Thank you to my fellow Curve denizens, my large adult sons, Jed Pizarro-Guevara and Nick Van Handel. Jed, you are an excellent senpai and friend. I'm grateful for everything you've taught me about stats, Philippine languages, anime, and more. Nick, I had a wonderful time sharing the Curve with you, even though we're both weirdo shut-ins, and you've made my dissertation markedly better more than once.

Thank you to the Es of J, who are, in addition to Tom, Steven Foley, Kelsey Kraus, and Deniz Rudin. Steven, working with you in z/lab was one of the best things I've ever gotten to do in school, and being friends with you is one of the best things I've ever gotten to do. Kelsey, you're the gold standard for Kelseys everywhere, and I will miss sowing Kelsey-related confusion with you. Deniz, you're one of my favorite universe coinhabitants. To honorary E of J Paul Willis, thank you for the impressively wide-ranging conversation topics and the fancy coffee.

Thank you to Maho Morimoto, for everything you did to build Nido, and for all the time we spent doing banda stuff. It all would've been much less fun and much more stressful without you. Thank you to the other early Nido builders: Andrew Hedding, Ben Eischens, and Jérémie Beauchamp.

Thank you to Stephanie Rich for the many good times, academic and not, in Santa Cruz and not, with ducks and not. Thank you to Lalitha Balachandran, for the fun hangs and the support. I'm very grateful for the times you reached out during this last little while.

Thank you to Margaret Kroll, for showing me that our path is righteous. To Jack Duff, for being a great fellow traveler in this odd pocket of the discourse world. To Netta Ben-Meir, for the sass and the sweetness, and for the recaps of bad TV shows. To Richard Bibbs, for cards and baseball. To Max Kaplan, for sports knowledge (non-baseball division), and for not being too bothered about that time at SFO. To Morwenna Hoeks, for being a great Maze partner, and for the good vibes you brought to Linghaus. To Hitomi Hirayama, for various short-but-lovely conversations, and for Japanese data.

To Bern Samko, for baseball and trivia. To Ben Meriçli, for unparalleled email game, interesting and amusing conversations, and late-night pizza on St. Marks. To Jeff Adler, for intentionally and unintentionally providing many laughs. To Andrew Angeles, Vishal Arvindam, Jenny Bellik, Nate Clair, Mansi Desai, Karl DeVries, Lisa Hofmann, Chelsea Miller, Anissa Zaitsu, and Erik Zyman, for making our little community so wonderful to inhabit.

Thank you to Alex Göbel, for illuminating discussions at conferences, and for being so patient and forgiving about my extremely slow emailing. To Adam Singerman, for funny times at late hours and top-notch animal pictures.

To my non-linguist friends (yes, I have those)—Willie, Brandon, Tristan, Siffert—thank you for the years of love and support.

Last, but certainly not least, to my family: thank you all, for all of it.

Chapter 1

Introduction

1.1 The question of coherence

Consider the pair of sentences in (1), and judge whether they seem to form a larger unit—a very short narrative, or perhaps a story fragment—or are merely two sentences that happen to be beside one another.

(1) Lora stood up. Yoshiko greeted her.

The most natural judgment is that these two sentences cohere somehow, to form a unified sequence of events. We infer a relationship between them: *First,* Lora stood up. *Then* Yoshiko greeted her.

On to another pair:

(2) Lora dropped a mug. It shattered.

These also seem to form a unit, linked not only by forward movement in time, as in (1), but also by causation: Lora dropped a mug. *As a result,* it shattered.

Now for a third pair:

(3) Lora fell. Yoshiko pushed her.

This pair is also a unit, linked by causality, but in a different temporal configuration than the pairs in (1) and (2). In this pair, the order of the sentences themselves seems to be the opposite of the order of the events they describe: Lora fell *because* Yoshiko *had* pushed her.

Finally, consider the pairs in (4) and (5).

(4) Lora stood up. Yoshiko adopted a guinea pig.

(5) Lora fell. Yoshiko has mixed feelings about artichokes.

These pairs do not seem to form coherent units—none of our preexisting world knowledge suggests that these pairs describe unified situations.

A central question in significant swaths of linguistic and psychological research is this: What makes a sequence of clauses or sentences seem like a coherent discourse? That is, what distinguishes the pairs of sentences in (1)-(3) from those in (4)-(5)? In the broadest strokes, the answer is that coherence is determined by inference processes stemming from our world knowledge, and influenced by linguistic information. If one can relatively readily infer that there is a plausible relationship between two described eventualities, then those descriptions can be interpreted as components of a single coherent discourse. In the past several decades, linguists and psychologists have made considerable progress towards understanding how exactly those inferences processes work. However, they have been working largely independently of one another, and an integrated perspective would open myriad new and productive lines of inquiry into understanding discourse coherence.

1.2 The goals of the dissertation

The goal of this dissertation is to lay some of the groundwork for a new kind of theory of discourse coherence. This involves bridging the gap between formal linguistic theories of discourse coherence and structure and psychological theories of discourse comprehension and its representation in memory. Although the major questions and empirical interests of these literatures overlap to a significant degree, there has been remarkably little dialogue between them. The current work only begins to explore the ways in which these literatures can challenge and inform one another. In broad strokes, the formal linguistics literature provides a variety of models of discourse structure that are quite finely articulated with respect to the effects of linguistic cues on the dynamics of discourse structure-building. A reasonable next step would be to put these theories to experimental tests—they are dynamic in the formal linguistic sense, but can they be operationalized as models of online discourse comprehension? And to what degree do they reflect the structure of discourse in memory? Psychological research on discourse provides a variety of theoretical and experimental tools with which to tackle these questions. In turn, the psychological literature on discourse structure provides a more domain-general framework for understanding discourse, in which linguistic input is viewed as a set of instructions for the construction of a mental model of a described situation. As linguists are keenly aware, though, the linguistic factors that may modulate discourse comprehension and representation are myriad, often individually complicated, and often interacting with one another in complex ways. It is valuable, then, to revisit and build on the experimental findings of psychological discourse theory with greater linguistic sensitivity.

This dissertation engages three main questions, given in (6)-(8).

(6) How does temporal iconicity influence discourse comprehension and discourse recall?

a. **Temporal iconicity** is when the order of events is matched by the order of the sentences that describe them. The sentence pairs in (1) and (2) are instances of temporal iconicity. The pair in (3) is an instance of **anti-**

iconicity—the order of the sentences is the opposite of the order in which the events they describe occurred.

(7) How does causal linking influence discourse comprehension and recall?

(8) What factors govern the comprehension of long-distance relationships in discourse, e.g., a pronoun in one sentence of a discourse that refers to an entity in a non-adjacent sentence?

1.3 Outline

Chapter 2 addresses the question posed in (6). It focuses on the role of temporal information (and intersentential temporal relations) in the online comprehension of discourse. Broadly, it addresses the assumption, common in both formal linguistic and psychological discourse coherence theories, that the default discourse relation is a non-causal progression, i.e., a relation with temporal iconicity, but no causal link, as in (1). This assumption is psychologically grounded—our only lived experience of time is as progression. However, these theories do not really address what kind of default this relation might be. Is it a first-choice default that comprehenders actively predict as they interpret a discourse? Or is it more of a last-resort default—only used once all other possibilities are eliminated? I present three self-paced reading experiments (Experiments 1A-1C) and one Maze task experiment (Experiment 1D) whose results suggest that, if non-causal progression is any sort of default, it is in the last-resort category, not the first-choice category.

Chapter 3 also takes up the question in (6), and advances to the question in (7). In this chapter, I turn to Situation Model Theory, a psychological theory of how coherent discourses are constructed and stored in memory. Through this lens, I focus on the roles of temporal and causal information in the structuring of discourse in memory. First, I use a probe recognition study in an attempt to adjudicate between conflicting previous findings about how 'story-time distance,' i.e., distance on a narrative's internal timeline, affects discourse structure (Experiment 2A). This experiment failed to replicate either of the previous results. Then, I investigate the role of causal information in discourse structuring in memory. Situation Model Theory considers causal linking quite important for discourse processing, yet has studied it relatively little (Experiments 2B and 2C). These studies, like Experiment 2A, fail to find story-time distance effects, but suggest that causal relations between sentences does indeed affect discourse structuring in memory.

Chapter 4 addresses the question in (8). To do so, I engage with a prominent linguistic theory of discourse coherence, Asher and Lascarides' Segmented Discourse Representation Theory, which unites dynamic semantics and a world knowledge-based inference system. I single out one part of this theory—the Right Frontier Constraint—for special scrutiny. I argue that it responsible for most of the clear, straightforwardly testable predictions that this theory makes. I address the question of whether the Right Frontier Constraint—which is the primary governor pronominal anaphora in discourse within the Segmented Discourse Representation Theory framework—must be encoded in the 'grammar' of discourse structure, or if it can actually be captured by a mixture of independent factors. Using the stops-making-sense experimental paradigm, I investigate the effects of animacy (Experiment 3A) and event structure (Experiment 3B) on pronominal anaphora in the context of Right Frontier Constraint violations. The results suggest that both animacy and event structure play a role in the comprehension of long-distance pronominal relationships in discourse. This challenges the necessity of the Right Frontier Constraint as a theory-internal, 'grammatical' condition on pronominal anaphora in discourse.

Chapter 5 concludes.

Chapter 2

'Defaults' in temporal interpretation

In this chapter, I focus on temporal interpretation in discourse. Understanding how cross-clausal temporal relations are formed is crucial to understanding discourse interpretation as a whole. Further, the temporal piece of the broader discourse puzzle is relatively simple for linguists to study, via the effects of overt linguistic cues, e.g., tense markers and temporal adverbials. Accordingly, significant swaths of formal discourse research are tightly focused on temporal relations. I provide an overview of some of this work in §2.1, and draw particular attention to how various theories of temporal interpretation in discourse characterize the progression temporal relation as a default. The conception of progression as default has deep non-linguistic roots—it reflects the way in which humans experience time. However, since humans can readily use language to describe things they cannot experience, it is worth scrutinizing the place of this default in linguistic theories. This leads to a discussion of the empirical adequacy of these theories, especially with regards to progression's opposite, backshift (1). The empirical challenge posed by backshift—particularly **unmarked backshift**, which is defined in (2)—was an important motivation for more pragmatically-developed theories of discourse. Among these theories, Lascarides and Asher's Segmented Discourse Representation Theory is especially prominent; I preview it in §2.2 and engage with it in depth in Chapter 4.

(1) **Backshift:** A backshift relation holds between two sentences (S1 S2) if the eventuality described in S2 is completed before the eventuality described in S1 begins.

(2) **Unmarked Backshift:** A backshift is unmarked if the temporal relation between S1 and S2 is not signaled by any of the following: tense, aspect, temporal adverbial, backshift-entailing connective (e.g., sentence-medial *because* in English).

Because backshift is a significant part of this examination of what sort of default progression is, in §2.3, I explore the empirical terrain of backshift through a corpus study of naturally-occurring backshifts. With formal and empirical foundations set, §2.4 reviews the small body of work investigating the processing of backshift, then presents a trio of self-paced reading experiments that tested the notion that progression is a first-choice default, and investigated some of the cues that comprehenders may use to determine discourse structure. Finally, §2.9 presents a Maze task experiment, a follow-up to one of the self-paced reading experiments, that provides a more fine-grained picture of how comprehenders integrate temporal information in discourse.

The guiding questions of the experiments in this chapter are these: To what degree can the processing of discourse structure be analogized to the processing of syntactic structure? Can/Do comprehenders make predictions about discourse structure, as they can—to a certain extent—in syntactic processing? To begin answering the second question, we need a clearer picture of what kind of 'default' progression is. Is it a first-choice default that a comprehender can actively predict during discourse processing, or is it a last-resort default? In the course of working this out, we also grapple with the question of what kinds of cues comprehenders use to interpret discourse structure, and when they use that information. That is, in discourse comprehension—as in syntactic comprehension—one thing we want to determine is whether comprehenders are interpreting discourse structure incrementally, as soon as they encounter cues to that structure, or whether they can/must delay their commitment to a particular discourse-level interpretation. We also want to know what cues, e.g., aspect markers like the English perfect 'had', comprehenders use to make such decisions, and the relative degrees of influence those cues seem to have. We also want to know how our nonlinguistic world knowledge influences discourse comprehension, as it figures heavily in linguistic theories such as Lascarides and Asher's, as well as theories of discourse from psychology, one of which I preview in §2.2 and engage with in detail in Chapter 3.

The results of my experiments suggest that progression is not a first-choice default, and that it may not even be a default in the sense of having the lowest baseline comprehension cost. However, there is some suggestion that the 'plausibility' of a given temporal relation, which is an umbrella for a range of causal and enablement inferences, does influence discourse processing. These takeaways are investigated further in Chapter 3.

2.1 The background of backshift

2.1.1 'Backshift' and temporal iconicity

The 'backshift' relation goes by various names across formal work in discourse interpretation, and seems to receive several different definitions. In this work, backshift is defined as in (3).

(3) **Backshift:** A backshift relation holds between two sentences (S1 S2) if the eventuality described in S2 is completed before the eventuality described in S1 begins.

Under this definition, S2s describing both eventive and stative/activity can be in a backshift relation with S1. For S2s describing stative and activity eventualities, this means that the state/activity cannot overlap with the eventuality in S1. In this respect, the definition in (3) diverges from that of, e.g., Asher & Lascarides (2003), for whom the eventuality in S2 must simply have started before the eventuality in S1, but is not required to have ended by that point.

Another important concept in this work is **temporal iconicity**—that our lived experience of time, i.e., forwards only, is reflected in the order of linguistic descriptions of eventualities. Progression relations are **iconic**, because the order of the sentences matches the order of the eventualities they describe (4). Backshift relations are **anti-iconic**, because the order of the sentences is the opposite of the order of the eventualities they describe (5).

(4) Maxine stood up. Joanna greeted her.

(5) Maxine fell. Joanna pushed her.

2.1.2 Sentence-internal accounts

The sentence-internal accounts are unified by their reliance on factors including tense and aspect markers, aktionsart, temporal adverbials, and linear order to determine temporal relations in discourse. The accounts reviewed here also share the view that tense is anaphoric, following Partee (1973). They also adopt a basically Reichenbachian (1947) view of speech time (SpT), reference time (RefT), and event time (EvT), and use semantic definitions of tense and aspect based on these temporal landmarks.[1]

Specifically, tense imposes requirements on the relationship between RefT and SpT, while aspect markers and aspectual adverbials impose requirements on the relationship between RefT and EvT. For example, the past tense means that the RefT must precede the SpT. As for aspect, the simple past (i.e., perfective aspect) means that the RefT and EvT may coincide, while the past perfect means that the EvT must precede the RefT.

This group is separable into two categories: those that assume or assert that unmarked backshifts are impossible, and which cannot account for them as originally formulated (Dowty, 1986; Hinrichs, 1986; Kamp and Reyle, 1993; a.o.); and those that acknowledge and allow unmarked backshifts (Webber, 1988; Kratzer, 1998).

2.1.2.1 Dowty 1986

Dowty proposes that temporal interpretation in narrative discourse depends on three factors: (i) and interval-semantic analysis of aktionsart[2], (ii) a principle governing the interpretation of consecutive sentences in a discourse, and (iii) 'a large dose of Gricean

[1]In the Reichenbachian system, the temporal landmarks are defined as in (i).

(i) a. Speech Time: The time at which the sentence is heard or read.
b. Reference Time: The time (frame) in which the eventualities in the sentence are set.
c. Event Time: The time at which the eventualities in the sentence occur/obtain.

[2]As developed in (Dowty, 1979).

conversational implicature and 'common sense' reasoning based on the hearer's knowledge of real work information' (41). His proposed general principle—the *temporal discourse interpretation principle* (TDIP)—is defined in (6).

(6) Given a sequence of sentences S_1, S_2..., S_n to be interpreted as a narrative discourse, the reference time of each sentence S_i (for i such that $1 < i \leq n$) is interpreted to be:

   a. a time consistent with the definite time adverbials in S_i, if there are any;

   b. otherwise, a time which immediately follows the reference time of the previous sentence S_{i-1} (45).

Per the TDIP, the default temporal relation is progression, which can be overridden only by a temporal adverbial. To allow a backshift interpretation between a past perfect-marked S2 and a simple past S1 while still abiding by the TDIP, Dowty says that the semantics of the past perfect are such that the EvT of S2 precedes its RefT. He also argues that unmarked backshift interpretations can arise with stative or activity S2s without violating the TDIP, by virtue of the interval-semantic properties of such sentences and of their attendant pragmatic inferences. Specifically, backshift is allowed in the sense that the beginning of a state or activity in S2 can be interpreted as preceding the event in S1, and the continuation of the state/activity is interpreted as overlapping with the event in S1. This overlap may occur if the pragmatically 'normal' assumption is that the state/activity began before the event in S1. Strictly speaking, the RefT still advances, but Dowty's 'large dose' of pragmatic reasoning can be deployed to achieve the right interpretation here.

When it comes to achievement and accomplishment S2s, though, unmarked backshifts do not seem to be possible. Given that achievement/accomplishment sentences lack the subinterval property, i.e., the property that if a sentence is true at interval *I* it is true at all subintervals of *I*, it must also lack the superinterval property. Then, the TDIP requires that an S1 and S2 that are both achievement/accomplishment sentences can only be interpreted as a progression, with no overlap, in the absence of an overt marker. Since Dowty admits that 'an enormous amount of real-world knowledge and expectation' is at work in discourse interpretation (52), one could imagine an extension of his proposal—or perhaps just a concretization of this pragmatic reasoning—that allows unmarked backshifts. For instance, one might say that 'common knowledge about the usual temporal relationships among events' (47) includes the knowledge that an implicit causality verb like 'scold' is often followed by an explanation of what led to the scolding. A backshift interpretation could thus be inferred, even in the absence of marking or any semantic requirement of backshift from an implicit causality verb.

As it stands, Dowty's proposal cannot automatically account for unmarked backshift interpretations. Various other accounts of temporal interpretation in discourse, including Hinrichs (1986) and Kamp and Reyle (1993), have similar gaps.

2.1.2.2 Kamp and Reyle 1993

Kamp & Reyle give a dynamic semantic account of temporal interpretation using Kamp's (1981) Discourse Representation Theory. On their account, temporal interpretation is driven entirely by tense, aktionsart, and the linear order of sentences. Following Partee (1973), tense introduces a temporal anaphor that establishes RefT by taking the RefT most recently established in the preceding discourse as its antecedent. The relationship between RefT and EvT is determined by aktionsart, which for them is syntactically encoded.

Kamp & Reyle's system, like Dowty's, allows overlap interpretations between an eventive S1 and a stative/activity S2, while disallowing unmarked backshifts with achievement/accomplishment S2s. As Asher and Lascarides (2003:62) observe, Kamp & Reyle predict that both discourses in (7) must be interpreted as progressions, though (7-b) is most naturally interpreted as a backshift.

(7) a. Maxine fell. Joanna helped her up.
    b. Maxine fell. Joanna pushed her.

The empirical coverage of Kamp & Reyle's account diverges from Dowty's when it comes to temporal adverbials—it does not allow a temporal adverbial alone to induce a backshift interpretation. They stipulate that sequences like (8) (their 5.68) are ungrammatical because the adverbial in S2 denotes an earlier time than does the adverbial in

S1, and temporal adverbials cannot change the unmarked relationship between RefT and EvT.

(8) The expedition arrived on the 25th of July 1941. It left on the 5th of May of that year.

They note that some speakers do find (8) acceptable, but propose only that the principle of temporal adverbial inertness be relaxed, not done away with altogether, to account for such speakers. As we will soon see, Asher and Lascarides combine the dynamic machinery of Discourse Representation Theory with a framework for pragmatic inference in discourse in a way that resolves this empirical inadequacy.

2.1.2.3 Webber 1988

Webber's account permits unmarked backshifts in limited cases. She proposes to treat tense as a discourse anaphor, in a way that allows unmarked backshifts under certain circumstances. She adopts the view developed in Moens and Steedman (1988) that events have a tripartite structure: an event is comprised of a preparatory phase, a culmination, and a consequent phase. As defined by Moens and Steedman (1988:16), the *culmination* is the point at which the event it completed—that is, it marks a change of state of the world from one in which the event has not been completed to one in which it has been completed. The pre-culmination state—the *preparatory process/phase*—

seems to include the event or sequence of events that cause the culmination (18). The post-culmination state—the *consequent state/phase*—includes some, but not necessarily all, of the events that are consequences of the event under consideration. Per Moens and Steedman, the consequent phase 'includes only those consequences that the speaker *views* as contingently related to other events that are under discussion, say by causing them or permitting them to occur' (16). It is not clear if they intended there to be a maximum number of contingent steps in either a preparatory phase or a consequent phase, or whether their contents are to be limited by anything other than the speaker's judgment of what counts as contingent. This gives the system the flexibility needed to allow unmarked backshift, but by way of burdening the pragmatic reasoning system in an unspecified manner. As we will see, Webber encounters some difficulties in defining the limits of this flexibility.

Webber proposes that the temporal anaphor of S2 links the RefT of S2 to the event in S1, thus establishing the EvT of S2. If the event in S2 is anaphoric to the consequent phase of the event in S1, a progression interpretation arises. If it is anaphoric to the culmination, we reach a simultaneous interpretation. If the event in S2 is anaphoric to the preparatory phase of S1's event, we get a backshift interpretation. In a departure from the accounts summarized above, Webber's account allows both overt temporal markers like aspect markers and temporal adverbials and pragmatic inference to determine which part of the event in S1 the temporal anaphor of S2 picks out as its referent. This allows for the backshift interpretation of (9).

(9) I told Frank about my meeting with Ira. We talked about ordering a Butterfly.

Webber argues that, in this case, it is the ambiguity of ‘we’ that makes both progression and backshift interpretations possible. If ‘we’ refers to the speaker and Frank, the event in S2 must be anaphoric to the consequent phase of the event in S1, so a progression interpretation arises. If ‘we’ refers to the speaker and Ira, the event in S2 is necessarily anaphoric to the preparatory phase of the event in S1, so only a backshift interpretation is possible in that case.[3]

The availability of backshift in this instance falls under Webber’s more general argument that unmarked backshift is possible if there is a way for the backshifted sentence to be interpreted as indirect speech or an embedded narrative. In (9), the S1 verb *tell (about)* may signal an embedded narrative. If this property is sufficient to allow a backshift interpretation, then a backshift interpretation should be possible for (9) no matter who the agent of S2 includes.

Webber also allows unmarked backshifts outside of indirect speech/embedded narrative contexts. For instance, she observes that a backshift interpretation is available for (10-a).

[3]It seems to me that backshift is possible for this example just so long as the agent of S2 includes the speaker or Ira, and excludes Frank.

(10) a. John went to the hospital. He took a taxi, because his car was in the shop.

b. John went to the hospital. He had taken a taxi, because his car was in the shop.

This is possible because *taking a taxi* can be interpreted as part of the preparatory phase of going to the hospital. Webber contends that this is contrasts with the sequences in (11), of which only the perfect-marked (11-b) can yield a backshift interpretation.

(11) a. John went to the hospital. #He broke his ankle, walking on a patch of ice.

b. John went to the hospital. He had broken his ankle, walking on a patch of ice.

This is where the underspecification of what may be included in a preparatory phase gets Webber's account into difficulties.[4]

Webber claims that *breaking one's ankle* cannot belong to the preparatory phase of *going to the hospital*, though there does not seem to be anything in the definition[5] of

[4]On top of the fuzzy definition of preparatory phases, Webber's underspecified criteria for backshifts further muddles this part of her discussion. In (10), it seems to me that *taking a taxi* is most naturally interpreted as a subpart of the act of *going to the hospital*. That is, it is not complete before the *going to the hospital* eventuality begins—it is not a backshift at all.

If *went to the hospital* is replaced with *arrived at the hospital*, as in (i), the contrast Webber intended does seem to appear.

(i) a. John arrived at the hospital. He took a taxi, because his car was in the shop.

b. John arrived at the hospital. #He broke his ankle, walking on a patch of ice.

[5]Such as it is—the *preparatory phase* does not seem to be explicitly defined by Moens and Steedman

the preparatory phase preventing it. In fact, it is not difficult to interpret the breaking of one's ankle as 'contingently related' to going to the hospital—it seems natural to interpret the ankle-breaking as the cause of the hospital-going, and therefore as part of the preparatory phase of going to the hospital. In this part of her discussion, Webber identified a need to rein in the possibilities of what can be part of a preparatory phase, but ultimately did not propose a solution. This is a similar challenge to the one Dowty faced in allowing that pragmatic reasoning plays a sizable role in intersentential temporal interpretation, but not delving far into the specific kinds of event knowledge and other world knowledge speakers use to carry out such reasoning.

2.2 Previewing Linguistic and psychological theories of discourse coherence

2.2.1 Segmented Discourse Representation Theory

Observing that (7-b), repeated here in (12), is readily interpreted as a backshift, despite the lack of explicit cues to that relation, Asher and Lascarides set out to build a theory that allows such an interpretation.

(12) Maxine fell. Joanna pushed her.

or Webber.

The result, Segmented Discourse Representation Theory (SDRT), does not focus directly and essentially exclusively on sentence-internal temporal markers, but rather posits a limited set of inferences that may hold between and within sentences in a discourse (Lascarides, 1992; Lascarides & Oberlander, 1993a, 1993b; Lascarides & Asher, 1993a). This set of inferences—or coherence relations—is founded on the notion that the default assumption guiding the interpretation of discourse is that discourses are coherent. The coherence relations encode reasoning based on both world knowledge and linguistic knowledge, which covers reasoning about intersentential temporal relations[6].

In SDRT, temporal relations between sentences are a consequence of the choice of coherence relation, not an independent phenomenon. Unmarked backshift is allowed, so long as the coherence relation between the relevant sentences allows that temporal relationship. This is not to say that sentence-internal factors are unimportant—markers like *had* and temporal adverbials constrain the set of relations that may be inferred between a pair of sentences. However, coherence relations provide a way around the undergeneration problem many sentence-internal accounts run into, while getting somewhat closer to the pragmatic reasoning system to which most of those accounts make reference.

Of their proposed coherence relations, Asher and Lascarides identify five as being

[6]The set of relations used in SDRT is similar to, but more articulated than, that proposed by Hobbs (1979, 1985)

most relevant to temporal relations. The interpretive effects of these relations are defined in (13). Narration and Result (13-a)-(13-b) entail a progression temporal relation between α and β; Elaboration (13-c) may give rise to a backshift interpretation; Background (13-d) is backshift-entailing under (what I understand to be) their definition of backshift, but not the one I assume here; and Explanation (13-e) is backshift-entailing for them and the current work.

(13) a. *Narration*(α, β): The event described in β is a consequence of (but not strictly speaking caused by) the event described in α.

b. *Result*(α, β): The event described in α caused the event or state described in β.

c. *Elaboration*(α, β): β's event is part of α's (perhaps by being in the preparatory phase, in the sense of Moens & Steedman 1988).

d. *Background*(α, β): The state described in β is the 'backdrop' or circumstances under which the event in α occurred (no causal connection but the event and state temporally overlap).

e. *Explanation*(α, β): The event described in β explains why α's event happened (perhaps by causing it).

Asher and Lascarides posit that Narration is the default coherence relation. This is in line with the (asserted or assumed) progression default in the previously described theories, with the added condition that the default is that there be no causal relation between the linked sentences.

2.2.2 Situation Model Theory

In psychology, the overarching question of discourse coherence is very similar to that of formal linguistics, but is concerned in particular with a discourse's representation in memory. Researchers in this domain address the following questions: (i) What are the components of a coherent representation of a discourse?; (ii) What factors allow (or disallow) comprehenders to build a coherent representation?; and (iii) How does the structure of that representation affect retrieval?

Situation Model Theory (SMT) is a prominent psychological theory of discourse coherence. It posits that discourses have three layers of representation in memory. First, there is the surface, or verbatim, representation of the linguistic input. Next, there is the 'textbase', or semantic representation. Finally, there is the situation model: a representation of the described situation, which is constructed based on our extralinguistic world knowledge in conjunction with the representations of the linguistic input itself, i.e., the first two layers. In SMT, linguistic content is viewed as 'processing instructions' for constructing a situation model (Zwaan and Radvansky 1998:162). Situation models are composed of a handful of 'dimensions,' which comprehenders monitor during discourse processing. Five dimensions are commonly proposed: temporal, causal, spatial, entity (characters and objects), and character goals. If there is a discontinuity in one of these dimensions, the situation model gets updated (creating a boundary roughly analogous to those between 'discourse segments' in linguistic discourse theories).

Most work in SMT tacitly assumes that progression/Narration is the default. This assumption is so tacit, in fact, that most work in SMT doesn't acknowledge the existence of backshift at all. More on this can be found in Chapter 3.

2.3 A corpus of backshifts in narrative

In this section, I outline several possible—and observed—ways to mark the start of a backshift.[7]

I also report on the results of a corpus survey of the Penn Discourse Treebank (2008). The Penn Discourse Treebank (PDTB) results are compared with results from

[7]The process of ending or exiting backshifts does not yet figure into the experimental part of this work, but I give a brief sketch here. It is not a straightforward process to find the endpoints of backshifts in the PDTB, so only examples from fiction are given at present. In these narratives, backshifts are often fairly lengthy flashbacks, but some are quite short, only spanning one or two sentences. Adverbials can signal the end of a backshift, often in combination with a paragraph break, as in ((i)a-b), but sometimes on their own ((i)c). A change from past perfect marking to simple past can also indicate the end of a backshift ((i)d).

(i) a. It **had** definitely been worthwhile. **Now** Carl **sat** in his easy chair in his study (Chiang:76).
 b. Because, despite all this wonderful technology, somehow...[he] **had** shaken off the tail they'd put on him after he'd resigned the last time. **So**, before they did anything else, Sma and it **had** to find the damn human first (Banks:790).
 c. Sma **had** arrived on the Xenophobe just after breakfast, by ship time. **When** she **awoke**, it was early afternoon (Banks:1034).
 d. The word **had** been pronounced from behind and to one side of her. She **saw** Sussepin's gaze falter and shift (Banks:347).

a corpus I have built, which focuses specifically on backshifts in narrative prose. The frequencies of different backshift marking strategies differ between the PDTB—which contains only expository prose—and the narrative prose corpus. Across these corpora, though, we have similar starting expectations, based on the assumption of default progression/Narration. First, backshifts should be overall less frequent than progressions. Second, overtly marked backshifts should be more frequent than unmarked backshifts (under the assumption that 'default' corresponds to 'unmarked'-ness).

While backshift occurs in past and present tense contexts alike, I focus here on past tense contexts only. The PDTB has yielded 124 examples of backshift in past tense contexts, and the narrative prose corpus currently contains 296 such examples. Three types of markers, listed in (14), were observed in these data, for both eventive and stative S2 verbs.

(14) a. Tense/aspect marking in S2: simple past or past perfect **had**

b. Temporal adverbial (e.g., **earlier**)

c. Discourse connective (e.g., **because**)

The possible combinations of these markers are laid out in Table 2.1.

Representative examples of these combinations of markers are presented in (15). The verbs that are in a backshift relationship are underlined, and the backshift markers are indicated in bold.

Table 2.1: Ways to mark backshift

	No Adv		Adv	
	No connective	Connective	No connective	Connective
Simple	ø	because	earlier	because + earlier
Perfect	had	because + had	earlier + had	because + earlier + had

(15) a. Simple past, no adverbial, no connective—Unmarked backshift

Virginia Vidaura—Envoy Corps trainer, later career criminal and sometime political activist. Something of a role model for me, though it was several decades since I'd last seen her. On a dozen different worlds, she crept into my mind unbidden, and I owed that ghost in my head my own life a dozen times over (Morgan 2005:10).

b. Simple past, overt adverbial, no connective

Mr. Sung couldn't be reached for comment. He **earlier** denied the allegations against him in the lawsuit by Norton and GE (PDTB)

c. Simple past, no adverbial, overt connective

He had Candlestick built **because** the Giants claimed they needed 10,000 parking spaces. (PDTB)

d. Past perfect, no adverbial, no connective

Secretary of State Baker expressed concern that Nicaraguan President Ortega may attempt to use alleged attacks by the U.S.backed Contra rebels as an excuse to scuttle elections scheduled for February. Ortega **had** threatened to end a 19-month-old ceasefire... (PDTB)

e. Past perfect, overt adverbial, no connective

My life had taken such a turn; this was most clear to me as I relaxed in this cowboy's house on an isolated ranch and watched John Wayne's swagger on the television screen before me. **Just months earlier**, living in LA, **I'd** found it difficult not to live by plans (R. Drummond 2011:99).

f. Past perfect, no adverbial, overt connective, stative

The rope that bound his hands behind the back of the chair was frayed, **because** he **had been** trying to cut through the bindings (Banks 2008:146).

Table 2.2 gives the raw counts of four types of backshift marking from the PDTB and narrative prose, as well as the proportions of the total examples (per prose type) that each row and column total represents. Examples with discourse connectives are not represented in Table 2.2 because they have not been the focus of the searches so far, so only a handful have been collected.[8]

Table 2.2: Raw counts of backshifts by prose type and marking type

PDTB (Expository)					Narrative		
	No Adv	Adv	Totals		No Adv	Adv	Totals
Simple	i) 34	ii) 39	73 (.59)		i) 11	ii) 6	17 (.06)
Perfect	iii) 37	iv) 14	51 (.41)		iii) 180	iv) 99	279 (.94)
Totals	71 (.57)	53 (.43)	124		191 (.65)	105 (.35)	296

With respect to adverbial marking, the expository and narrative data are distributed

[8] This exclusion is due to the focus of the first stage of the experimental part of this project, which only includes aspect and temporal adverbial marking.

in approximately the same way. Looking at the column totals (in blue), we see that for both prose types, backshifts with no adverbial marker outnumber those with adverbial markers by roughly 60-40. With respect to aspectual marking, the prose types differ quite a bit from each other. Looking now at the row totals (in grey), we see that the expository data comprise about 60% simple past backshifts, and 40% past perfect (i.e., *had*-marked) backshifts. In contrast, the narrative data comprise only 6% simple past backshifts.

The distributions of individual marking combinations also differ across prose types. The proportions of each marking combination are in Table 2.3.

Table 2.3: Proportions of backshifts by marking type

PDTB (Expository)				Narrative		
	No Adv	Adv			No Adv	Adv
Simple	i) .27	ii) .31		Simple	i) .04	ii) .02
Perfect	iii) .3	iv) .11		Perfect	iii) .61	iv) .33
	n=124				n=296	

For both prose types, about 60% of the examples are single-marked backshifts—i.e., just an adverbial or just *had*—though the proportions of the specific marking types differ. When it comes to double-marked backshifts, i.e., *had* + *earlier* cases, only 11% of the PDTB backshifts had such marking, but a third of the narrative prose cases are double-marked. Perhaps most interesting, though, is the comparison of unmarked backshifts (the (i) cells). Of the PDTB backshifts, 27% are unmarked, while only 4% of the narrative backshifts are unmarked.

If it is the case that progression is the default temporal relation in narrative discourse, the paucity of unmarked backshifts among the narrative examples does not seem so surprising. Appealing to neo-Gricean principles can help us sharpen this intuition. Horn (1984):22 argues that a general pattern in cooperative discourse is that ”the use of a marked expression when a corresponding unmarked expression is available tends to be interpreted as conveying a marked message (one which the unmarked alternative would not or could not have conveyed).” Assuming progression is the default, and that simple past with no adverbial is the unmarked form, one would expect the unmarked form to be associated with the progression interpretation, and thus for it to be associated with the backshift interpretation only rarely. Amongst the narrative prose examples, this is the case. For the expository examples, though it is not. In the PDTB cases, there are nearly as many unmarked backshifts as adverbial-marked or *had*-marked backshifts. Specific features of the PDTB cases might suffice to explain this difference, though we cannot yet rule out the possibility that some more fundamental difference between expository and narrative discourse is partly responsible.

Many of the backshifts collected from the PDTB—and most of the unmarked backshifts—are templatic. That is, they occur in a handful of types of news items. For example, nearly two-thirds of the unmarked backshifts are ”new executive” announcements (16-a) or ‘ticker’ announcements (16-b).

(16) a. Denis C. Smith **was named** president and chief operating officer of Big

Corporation. Mr. Smith **was** senior executive vice president...

b. Bank X **gained** 2/14 to 72 7/8. The Bank **agreed** to be acquired in a merger...

The templatic nature of these data gives us a plausible way to account for the frequency of unmarked backshifts. Whatever the general pragmatic principles that militate against unmarked backshifts, these can be overridden in the limited, specialized contexts in which most of these cases appear in the PDTB. It does not give us a straightforward way to account for the scarcity of adverbial-marked backshifts in the narrative examples—it is not obvious why this method of marking the marked interpretation is dispreferred compared to *had*-marking or double marking.

Finally, it is worth noting that backshift is fairly uncommon in general, regardless of prose type. The PDTB examples were drawn from approximately 2000 passages with backshift-compatible annotations, and the narrative examples from twelve books.

2.4 The processing of backshift

There is a small body of psycholinguistic work on the comprehension of temporal relations in discourse. Here, I summarize Dickey's (2001) work, which serves as the jumping-off point for my first three experiments (1A-1C).

Dickey draws on Bestgen and Vonk's (1995, 2000) findings on the role of temporal adverbials in discourse comprehension. In a self-paced reading task, Bestgen and Vonk (2000) found that sentence-initial temporal adverbials such as ‘around eleven o’clock’ ameliorated the *boundary effect*. The boundary effect is a phenomenon in which the reading times for sentences that introduce a new episode/topic (where topic is quite hazily defined) are higher than for intraepisodic/nontopic sentences (Haberlandt, Berian, & Sandson, 1980; Lorch, Lorch, & Matthews, 1985). For instance, in the example discourses in (17) and (18), from Bestgen & Vonk’s (2000) experiments, the bolded sentence was read more slowly in (18) than in (17), since in (18) it introduces a new episode/topic in the discourse[9].

(17) This Monday, I got up very late. I had a full breakfast. I decided to go for a trip in the country. I dressed myself warmly.

I took my moped from the garage.

The front tire was completely flat...

(18) I went into the kitchen to prepare the dinner. I peeled the potatoes. I put the roast in a saucepan.

I took my moped from the garage.

The front tire was completely flat...

The non-temporal adverbials Bestgen and Vonk tested, e.g., ‘as usual,’ did not have

[9] Bestgen & Vonk’s experiments were conducted in French, but only the English translation is shown here because the boundary effect is expected to obtain regardless of the language.

the same effect, nor did post-verbal temporal adverbials. They consequently argued that sentence-initial temporal adverbials triggered the construction of a new discourse segment.

Dickey used this interpretation to link a coherence view of discourse structure (following Lascarides & Asher 1993, essentially) with his version of the parallel processor. He argued that, assuming coherence relations cannot exist between discourse segments, a sentence-initial temporal adverbial should block coherence relations—and therefore also temporal relations—from being formed between the sentence that the adverbial is part of and the preceding sentence. He then sought to test this by investigating the processing of backshifts—an extension of sorts of Bestgen & Vonk, who only tested progressions.

Dickey argued that, under this parallel/discourse-driven model, backshifts signaled by sentence-final (or at least post-verbal) adverbials should be more costly to comprehend than progressions signaled by final adverbials. This follows under the assumption that progression is the default interpretation: a strongly incremental parser should commit to a progression interpretation in the absence of overt, fully disambiguating cues before it encounters a final adverbial. When it does encounter a final backshift-signaling adverbial, a 'discourse garden-path effect' should arise. In contrast, backshifts signaled by initial adverbials should not be more costly than progressions signaled the same way. Since initial temporal adverbials sever coherence/temporal relations on this view, no garden-pathing can occur.

To test this prediction, Dickey ran a pair of self-paced reading studies, both with a 2x2 design crossing adverbial position (Initial, Final) with coherence relation (Background, Narration)[10]. His first experiment (N=44) had 14 items of the form in (19).

(19) Last summer a senior aide in the Senator's press office was accused of sexually harassing a female staffer. |

He denied | the claims | but resigned | from the staff | anyway. |

a. Initial AdvP Narration: At his next job | he took | advantage of | another staff member.

b. Initial AdvP Background: At his previous job | he took | advantage of | another staff member.

c. Final AdvP Narration: He took | advantage of | another staff member | at his next job.

d. Final AdvP Background: He took | advantage of | another staff member | at his previous job.

Dickey found that the AdvPs (e.g., 'at his previous job') of Final Backgrounds (19-d) were slower than those of Final Narrations (19-c), but that the AdvPs of Initial Back-

[10]Dickey uses relations as defined by Asher and Lascarides. Background entails a backshift temporal relation without a causal relation between S1 and S2. Narration entails progression without a causal relation between S1 and S2.

grounds (19-b) were faster than those of Initial Narrations (19-a) (for corrected RTs). However, the fact that the critical region (AdvP) in the Final conditions coincided with the end of a sentence meant that any effect of Coherence Relation in those conditions would not be separable from general sentence wrap-up effects (Just & Carpenter, 1980).

In hopes of distinguishing possible effects of Coherence Relation in the Final conditions from sentence wrap-up effects, in his Experiment 2, Dickey added a spillover region after the AdvP (N=60; 24 items). His Experiment 2 items were of the form in (20).

(20) Last summer | a senior aide | in the Senator's | press office | was accused | of sexually harassing | a female staffer. |
He denied | the claims | but resigned | from the staff | anyway. |

a. Initial (AdvP) Narration: At his next job | on Capitol Hill | he took | advantage of | another staff member.

b. Initial (AdvP) Background: At his previous job | on Capitol Hill | he took | advantage of | another staff member.

c. Final (AdvP) Narration: He took | advantage of | another staff member | at his next job | on Capitol Hill.

d. Final (AdvP) Background: He took | advantage of | another staff member | at his previous job | on Capitol Hill.

Here, Dickey found that Narrations were slower than Backgrounds at AdvP, regardless of adverbial position. He put this down to systematic length differences between his Narration and Background AdvPs. Crucially, he found that, in the Spillover regions, Backgrounds were slower than Narrations regardless of position. He took this to show that the temporal relationship the Background adverbial signaled was 'non-default...dispreferred and costly' (p.109). He further argued that there was no evidence of garden-pathing in the Final conditions, since the Spillover effect obtained in Initial position as well.

2.5 Experiment 1A: AdvP Position x Temporal Relation

2.5.1 Design and Predictions

This experiment followed the manipulation made in Dickey's experiments: it had a 2x2 design, with the factors AdvP Position {Initial, Final} and Temporal Relation {Backshift, Progression}. However, it was not designed to be an exact replication attempt. While Dickey's stimuli had three sentences each, mine had two. This was done under the assumption that comprehenders attempt to establish coherence even when presented with only two sentences/clauses. This assumption seems to be intuitively grounded, and to be used to a degree in coherence theory (consider, for instance, *Max-*

ine fell. Joanna pushed her.). However, I will reexamine this assumption in §2.7.6. Participants were instructed that they would be reading very brief stories—the information that they would be presented with presumably coherent narratives was expected to be enough to induce coherence establishment.

Like Dickey's experiments, the Experiment 1 stimuli were presented in phrase-by-phrase self-paced reading. The regions are mapped out in Table **??**. The second change from Dickey's design was the addition of two more spillover phrases after the AdvP. This was to further separate possible effects of temporal relation appearing at the disambiguating AdvP from the end of the sentence in the Final AdvP conditions, as well as to further separate the AdvP from the SubVerb and V+1 regions in the Initial AdvP conditions.

Table 2.4: Self-paced reading regions in Experiment 1.

Sent1		
Lianne poked Eric.		
AdvP	**Spill1**	**Spill2**
A moment earlier,	in his seat	near the back
Spill3	**SV**	**V+1**
of the lecture hall,	he pinched	her on the arm.

On a view in which progression is a strong default, i.e., it is the temporal relation that comprehenders actively predict, we expect to see that progression has an advantage over backshift in self-paced reading measures. That is, we expect to see a kind of **iconicity effect**, in this case, a **progression advantage effect**. If it is the case that the comprehender posits progression immediately after reading sentence 1, encountering a backshift AdvP (e.g., *a moment earlier*) at any point in sentence 2 should incur a reanalysis penalty. Thus, we expect the Backshift conditions to be read more slowly than the Progression conditions as early as the AdvP region, whether the AdvP is sentence-initial or sentence-final. Since self-paced reading effects often emerge only in spillover, we also expect to see the progression advantage effect somewhere in the Spill1-Spill3 regions.

2.5.2 Materials

Thirty-six experimental stimuli were constructed. All stimuli were two sentences long. The first sentence was constant across all four experimental conditions. All stimuli had 'Low' backshift plausibility, as determined in a separate norming study. (See Appendix A for details of the norming study.) A sample item is in (21); the complete materials are in Appendix B.1. The pipes (|) indicate the boundaries of the phrase-by-phrase presentation.

(21) **S1:** Lianne poked Eric. |

a. **S2** (Init, Back): A moment earlier, | in his seat | near the back | of the lecture hall, | he pinched | her on the arm.

b. **S2** (Init, Prog): A moment later, | in his seat | near the back | of the lecture hall, | he pinched | her on the arm.

c. **S2** (Fin, Back): He pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

d. **S2** (Fin, Prog): He pinched | her on the arm | a moment later, | in his seat | near the back | of the lecture hall.

The experimental stimuli were combined with 72 fillers. The fillers were list-invariant, consisted of two sentences each, and had the same number of presentation regions (7) as the experimental stimuli. All 108 stimuli were followed by a polar comprehension question.

Four experimental items were excluded from the analysis, for a total of 32 experimental items included in the final analysis. One was excluded because of a typographical error, three because at least two-thirds of participants gave incorrect responses to the comprehension question.

2.5.3 Participants

All participants provided informed consent, and were asked to report their age and basic linguistic background information. Participants were recruited through Mechanical Turk, and were paid $4 for completing the experiment. The experiment took approximately 40 minutes to complete. Participation was restricted to workers with IP addresses in the United States who had not participated in any of the other studies.

Sixty self-identified native English speakers were recruited. Twenty participants were excluded from the analysis for low comprehension question accuracy (¡80%). A further four participants were excluded because they did not seem to have completed a majority of the task as intended.[11]

2.5.4 Procedure

The 36 experimental items were distributed into four Latin Square lists. The order of presentation was randomized. The experiment was administered online via IbexFarm (Drummond, 2016). Participants read items in phrase-by-phrase, centered self-paced reading format, through which they advanced by pressing the spacebar. On the following screen, they responded to the comprehension question by clicking their response or by using the 'f' and 'j' keys. Participants were allowed to take as much time as they liked to read through the items, but were allowed a maximum of 5 seconds to answer the comprehension question. Participants were informed of the time limit, and instructed to respond as quickly and accurately as possible. When participants timed out on a comprehension question, they were automatically advanced to the next item. Before the task proper, participants were given written instructions, then completed three practice trials.

[11]More than half of these participants' raw reading times were extremely short given the region length, and many of those low reading times were identical, suggesting that the participants had simply held the spacebar down.

2.5.5 Results

Data were analyzed in R (2013). Reading times were log-transformed, then residualized using a linear mixed effects model (D. Bates, Maechler, Bolker, Walker, et al., 2014), into which region length and region position within an item were entered as fixed effects. Random intercepts were estimated for subjects and items. For the main analyses, only data from items for which the comprehension questions were answered correctly were used. The residualized data were analyzed using linear mixed effects models with AdvP Position and Temporal Relation, and their interaction, as fixed effects. Random intercepts and random slopes for the fixed effect parameters were estimated for subjects and items in all models unless otherwise noted. Sum coding was used (Final AdvP = +, Backshift = +). Simple effects were also analyzed separately.

The mean residualized RTs for each region in Initial AdvP conditions are presented in Figure 2.1. The mean residualized RTs for Final AdvP conditions are in Figure 2.2.

2.5.5.1 Temporal Relation

Analysis with sum coding revealed significant effects of Temporal Relation at the AdvP region: Progressions were read more slowly than Backshifts (β=-.04 (.02), t=-2.06). Pairwise comparisons showed that Final Progressions were slower than Final Backshifts at AdvP (β=.06 (.02), t=2.53). Initial Progressions and Backshifts did not differ at this region (t=.52). At Spillover1, the analysis with sum coding showed that

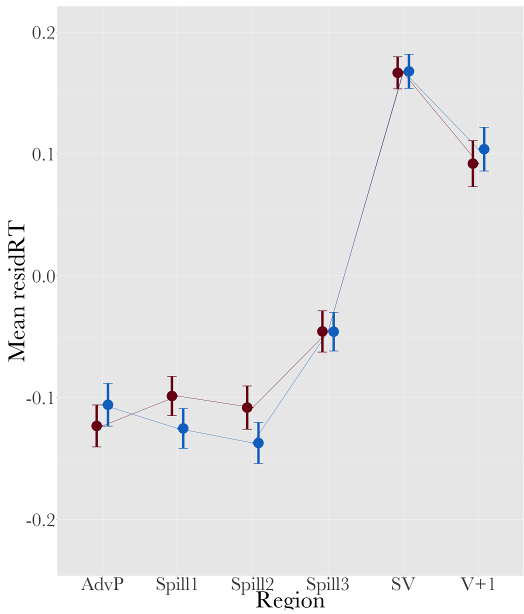


Figure 2.1: Mean residual reading times for Initial conditions in Experiment 1A. Error bars show standard error of the mean.

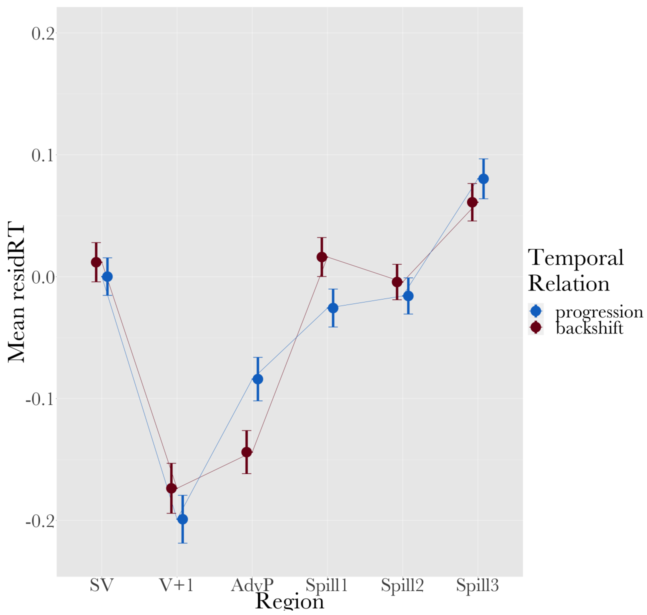


Figure 2.2: Mean residual reading times for Final conditions in Experiment 1A. Error bars show standard error of the mean.

Backshifts were read more slowly than Progressions (β=.04 (.02), t=2.26). In pairwise analyses, however, no significant differences obtained. There was a trend in the same direction in the comparison between Final Backshifts and Final Progressions (t=-1.87; reference level Final Backshift). Since the pairwise results are from maximal models, while the sum coded results are not, I will take the pairwise results to be the more reliable ones. Furthermore, the differences observed in the sum-coded analysis would not be significant under correction for multiple comparisons. The results of the analysis with sum coding are summarized in Table 2.5. The results of the pairwise analyses are summarized in Tables 2.6-4.

2.5.5.2 AdvP Position

Both sum-coded and pairwise analyses revealed significant effects of AdvP Position at all regions except the AdvP itself. Specifically, all three Spillover regions (i.e., spillover from AdvP) were read more slowly in the Final AdvP conditions they were in the Initial AdvP conditions. That is, when the Spillover regions were read more slowly when they were the last regions in the stimuli. Similarly, the SubVerb and V+1 regions were read more slowly when they were the last regions in the stimuli, i.e., in the Initial AdvP conditions.

Table 2.5: Summary of linear mixed effects models in Experiment 1A. Sum coded. Bold cells indicate significant coefficients. Temporal Relation was coded Backshift +, Progression -. AdvP Pos was coded Initial +, Final -.

	AdvP		Spill.1*		Spill.2	
	β	t	β	t	β	t
Temp Rel	**-.04 (.02)**	**-2.06**	**.04** (.02)	**2.26**	.02 (.02)	1.11
AdvP Pos	.003 (.03)	.13	**.1** (.02)	**6.4**	**.12** (.02)	**6.39**
TR:AP	-.05 (.04)	-1.3	.01 (.03)	.37	-.03 (.04)	-.72
	Spill.3**		SubVerb		V+1***	
	β	t	β	t	β	t
Temp Rel	.01 (.02)	.31	.004 (.02)	.29	.004 (.02)	.29
AdvP Pos	**.12** (.02)	**5.94**	**-.16** (.02)	**-8.68**	**-.16** (.02)	**-8.68**
TR:AP	-.01 (.03)	-.31	.01 (.03)	.33	.01 (.03)	.33

*Random slopes for subjects, full items **Random intercepts for items, full sub
***Random slopes for items, full sub

Table 2.6: Summary of linear mixed effects models for pairwise comparisons in Experiment 1A. Reference level Initial Backshift. Bold cells indicate significant coefficients.

	AdvP		Spill.1*		Spill.2**	
	β	t	β	t	β	t
Init Prog	.01 (.03)	.52	-.03 (.02)	-1.42	-.03 (.03)	-1.23
Fin Back	-.02 (.03)	-.67	**.11 (.02)**	**4.73**	**.1 (.02)**	**4.21**
TR:AdvP Pos	.05 (.04)	1.3	-.01 (.03)	-.37	.03 (.04)	.75
	Spill.3***		SubVerb		V+1****	
	β	t	β	t	β	t
Init Prog	.001 (.02)	.05	.001 (.02)	.03	.01 (.03)	.49
Fin Back	**.11 (.02)**	**5.17**	**-.16 (.02)**	**-6.66**	-.26 (.03)	-7.98
TR:AdvP Pos	.01 (.03)	.45	-.01 (.03)	-.33	-.04 (.04)	-.9

*Uncorrelated items, full sub **Random intercepts for items, full sub
***Full items, random intercepts for sub ****Uncorrelated items, full sub

Table 2.7: Summary of linear mixed effects models for pairwise comparisons in Experiment 1A. Reference level Final Progression. Bold cells indicate significant coefficients.

	SubVerb		V+1*		AdvP**	
	β	t	β	t	β	t
Fin Back	-.01 (.02)	-.46	-.03 (.03)	-.79	**.06 (.02)**	**2.53**
Init Back	**.16 (.02)**	**6.66**	**.26 (.03)**	**7.98**	.02 (.03)	.67
TR:AdvP Pos	.01 (.03)	.33	.04 (.04)	.9	-.05 (.04)	-1.3
	Spill.1		Spill.2***		Spill.3	
	β	t	β	t	β	t
Fin Back	-.04 (.02)	-1.87	-.01 (.02)	-.31	.01 (.02)	.46
Init Back	**-.11 (.02)**	**-4.73**	**.2 (.03)**	**-4.08**	**-.12 (.02)**	**-4.78**
TR:AdvP Pos	.01 (.03)	.37	-.03 (.04)	-.72	-.01 (.03)	-.31

*Random slopes for items, full sub **Uncorrelated items, full sub
***Full items, random slopes for sub

2.5.6 Discussion

This experiment did not replicate the results of Dickey's study. Dickey found that, at the spillover region after the AdvP, backshifts were read more slowly than progressions, whether the AdvP was Initial or Final. In the current experiment, there was no difference between Initial Backshifts and Initial Progressions at any region. In the Final conditions, there was an unexpected difference between Progressions and Backshifts—the Progressions were read more slowly. However, this effect was quite small, and likely would not survive correction for multiple comparisons. And, while the Final condition Spillover1 results show a trend towards a progression advantage, the effect did not reach significance. The AdvP position effects are, at least in part, consistent with a general wrap-up effect (Just & Carpenter, 1980).

While Dickey's findings suggested that backshifts were generally more costly to process than progressions, the results of this experiment are not straightforwardly compatible with that view. This casts some doubt on the conception of progression as a strong default, which has reflexes in on-line processing. Since all the items had Low backshift plausibility, and since no other highly valid cues for backshift interpretations were present, we might have expected that the default progression interpretation should have been privileged over backshift interpretations in all of them. On a coherence-centric view, one might expect that the comprehender assumes a progression/Narration interpretation by default, and would not change this assumption since there are no High backshift plausibility items to raise the baseline expectation of a non-default relation. It could also be the case, though, that the comprehenders in this experiment learned that there were no (highly) valid backshift cues, such as the past perfect 'had', in the experimental context, yet backshifts still occurred. This might have boosted their baseline expectation of backshift. Either way, if progression's default nature really does give it a baseline comprehension advantage over other temporal relations, it does not seem to be a very large or strong advantage.

The lack of temporal relation effects may also be due in part to the differences between Dickey's materials and the current ones. Dickey's items were three sentences, while the current items were two sentences. The first sentences of Dickey's items were intended to establish that the discourse is set in the past, with the experimental manipulations occurring between the second and third sentences. It is possible that this additional sentence was enough to create a stronger expectation of progression than my two-sentence items. That is, progression seems to be a default of *narrative* discourse, and perhaps the comprehender must first establish that they are dealing with a narrative discourse before they begin making the predictions appropriate to that type of discourse. Since establishing that a discourse is a narrative itself requires at least two sentences, these stimuli may have been too short for a progression advantage effect to emerge.

It is reasonable to wonder whether the lack of difference between backshifts and progressions overall in this experiment is due to comprehenders becoming more aware of backshifts occurring in the experiment, and thus becoming faster at processing them as the experiment progressed. An informal order analysis suggests that this is not the case. The mean reading time for the first eight backshift stimuli was 5305.95 ms, while the mean reading time for the last eight backshift stimuli was 5304.89 ms. Since participants do not seem to have become markedly faster at reading backshifts by the end of the experiment, we have some further support for the position that backshift and progression do not have very different baseline processing costs.

2.6 Experiment 1B: Plausibility x Temporal Relation

2.6.1 Design and Predictions

Experiment 1B investigates how pragmatic information—specifically, the plausibility of a backshift interpretation vs. a progression—modulates the comprehension of temporal relations. It had a 2x2 design, with the factors Plausibility of Backshift {High, Low} and Temporal Relation {Backshift, Progression}. 'Plausibility' here is an umbrella term for pragmatic reasoning based on various semantic and world-knowledge factors, from implicit causality—a lexical-semantic property of a verb—up to general world knowledge about the order a certain pair of events is most likely to occur in. The Plausibility factor was intended to capture relationships between the first sentence and the SV(O) of the second sentence. So, to make it easier to separate Plausibility effects from Temporal Relation effects, all items had Final AdvPs. That way, the disambiguating information in the temporal AdvPs would not intervene between S1 and the SV(O) of S2.

Under the assumptions (also made for Experiment 1A) that a two-sentence discourse is sufficient to coax out progression advantage effects, and that progression is a strong default, we make two main predictions for this experiment. First, in the Low conditions, which are the same as the Final AdvP conditions in Experiment 1A, progressions should be read more quickly than backshifts at the AdvP and Spillover regions. Second, if it is the case that Plausibility influences the comprehension cost of backshift, we expect an interaction: the difference between backshift and progression in the High conditions should be smaller than in the Low conditions, if not leveled altogether.

2.6.2 Materials

Thirty-six experimental stimuli were combined with the same 72 fillers used in Experiment 1A. Comprehension questions were constructed and presented in the same way as in Experiment 1A. Four experimental stimuli were excluded from the analysis, for a total of 32 items in the final analysis. One was excluded due to typographical error, three because at least 2/3 of participants responded incorrectly to the comprehension question.

A sample item is in (22); the complete materials are in Appendix B.2.

(22) **S1:** Lianne poked Eric. |

a. **S2** (High, Back): He fell asleep | in biology class | a moment earlier, | in his seat | near the back | of the lecture hall.

b. **S2** (High, Prog): He fell asleep | in biology class | a moment later, | in his seat | near the back | of the lecture hall.

c. **S2** (Low, Back): He pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

d. **S2** (Low, Prog): He pinched | her on the arm | a moment later, | in his seat | near the back | of the lecture hall.

2.6.3 Participants

Participant recruitment, payment, and experiment duration were the same as for Experiment 1A (§2.5.3). Thirteen participants were excluded from the analysis on the basis of comprehension question performance, and four were excluded due to apparent spacebar-holding. The remaining 43 participants are included in the analysis.

2.6.4 Procedure

The procedure was the same as in Experiment 1A (§2.5.4).

2.6.5 Results

The data were analyzed via the same procedures used for Experiment 1A (§2.5.5).

The mean residualized RTs for each condition are presented in Figure 2.3. The analyses with sum coding revealed a significant effect of Plausibility at the V+1 region: Low conditions were read more slowly than High conditions (β=-.05 (.02), t=-2.74). Pairwise comparisons at V+1 showed that Low Backshifts were read more slowly than High Backshifts (β=.08 (.02), t=2.95 with reference level High Backshift).

At the AdvP region, sum coding analysis showed the same effect of Plausibility (β=-.06 (.02), t=-3.37). There was also a significant effect of Temporal Relation: Progressions were read more slowly than Backshifts (β=-.04 (.02), t=-2.22). There was no interaction between the two factors. Pairwise comparisons revealed that Low Backshifts were read more slowly than High Backshifts (β=.08 (.03), t=3.23 with reference level High Backshift). High Progressions were also read more slowly than High Backshifts (β=.06 (.03), t=2.06 with reference level High Backshift).

At Spillover2, there was an interaction (β=-.07 (.03), t=-2.07 with sum coding).

The results of the analyses with sum coding are summarized in Table 2.8. The results of the pairwise comparisons are summarized in Tables 2.9, 2.10, and 2.11.

Table 2.8: Summary of linear mixed effects models in Experiment 1B. Sum coded. Bold cells indicate significant coefficients. Temporal Relation was coded Backshift +, Progression -. Plausibility was coded High +, Low -.

	SubVerb*		V+1*		AdvP**	
	β	t	β	t	β	t
TR	.02 (.02)	1.21	-.02 (.02)	-1.03	**-.04 (.02)**	**-2.22**
Pl	-.03 (.02)	-1.39	**-.05 (.02)**	**-2.74**	**-.06 (.02)**	**-3.37**
TR:Pl	-.01 (.03)	-.31	-.05 (.04)	-1.14	-.05 (.04)	-1.43
	Spill.1		Spill.2***		Spill.3	
	β	t	β	t	β	t
TR	-.01 (.02)	-.72	-.02 (.02)	-1.28	-.001 (.01)	-.14
Pl	-.001 (.02)	-.04	-.01 (.02)	-.38	.01 (.01)	.99
TR:Pl	-.04 (.03)	-1.22	**-.07 (.03)**	**-2.07**	-.01 (.03)	-.22

*Full items, random slopes for subjects **Random intercepts for items, full subj
***Random slopes for items and subjects

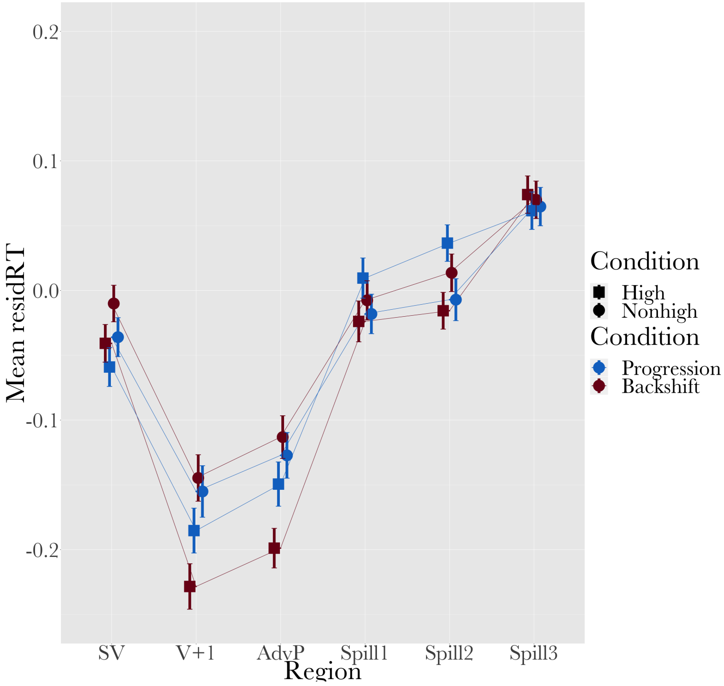


Figure 2.3: Mean residual reading times for all regions in Experiment 1B. Error bars show standard error of the mean.

Table 2.9: Summary of linear mixed effects models for pairwise comparisons in Experiment 1B. Reference level Low Progression. Bold cells indicate significant coefficients.

	SubVerb		V+1*		AdvP	
	β	t	β	t	β	t
Low Back	.02 (.02)	1.07	.01 (.03)	.18	.01 (.03)	.33
High prog	-.02 (.02)	-.97	-.03 (.03)	-.98	-.01 (.03)	-.44
TR:Plaus	-.01 (.03)	-.31	-.05 (.04)	-1.22	-.07 (.04)	-1.59
	Spill.1		Spill.2		Spill.3	
	β	t	β	t	β	t
Low Back	.01 (.02)	.41	.01 (.02)	.59	.001 (.02)	.06
High prog	.02 (.02)	.79	.03 (.03)	1.14	.02 (.02)	.88
TR:Plaus	-.04 (.03)	-1.22	**-.07 (.03)**	**-2.07**	-.01 (.03)	-.22

*Random slopes for items, full subjects

Table 2.10: Summary of linear mixed effects models for pairwise comparisons in Experiment 1B. Reference level High Progression. Bold cells indicate significant coefficients.

	SubVerb		V+1		AdvP*	
	β	t	β	t	β	t
High Back	.01 (.02)	.67	-.04 (.03)	-1.68	**-.06 (.03)**	**-2.16**
Low Prog	.02 (.02)	.97	.03 (.03)	1.01	.01 (.03)	.47
TR:Plaus	.01 (.03)	.31	.05 (.04)	1.14	.07 (.04)	1.78
	Spill.1		Spill.2		Spill.3	
	β	t	β	t	β	t
High Back	-.03 (.02)	-1.3	**-.06 (.02)**	**-2.34**	-.005 (.02)	-.25
Low Prog	-.02 (.02)	-.79	-.03 (.03)	-1.14	-.02 (.02)	-.88
TR:Plaus	.03 (.03)	1.22	**.07 (.03)**	**2.07**	.01 (.03)	.22

*Full items, random intercepts for subjects

Table 2.11: Summary of linear mixed effects models for pairwise comparisons in Experiment 1B. Reference level High Backshift. Bold cells indicate significant coefficients.

	SubVerb*		V+1*		AdvP	
	β	t	β	t	β	t
Low Back	.03 (.03)	1.23	**.08 (.02)**	**2.95**	**.08 (.03)**	**3.23**
High Prog	-.01 (.02)	-.67	.04 (.02)	1.68	**.06 (.03)**	**2.06**
TR:Plaus	-.01 (.03)	-.31	-.05 (.04)	-1.14	-.06 (.04)	-1.59
	Spill.1		Spill.2**		Spill.3***	
	β	t	β	t	β	t
Low Back	.02 (.02)	.93	.04 (.02)	1.87	-.01 (.02)	-.56
High Prog	.03 (.02)	1.3	**.06 (.02)**	**2.34**	.005 (.02)	.27
TR:Plaus	-.04 (.03)	-1.22	**-.07 (.03)**	**-2.07**	-.005 (.03)	-.18

*Random slopes for items and sub **Full items, random slopes for sub
***Random intercepts for items and sub

2.6.6 Discussion

The effect of Plausibility at V+1 suggests that the discourse-level cues it provides can be integrated fairly quickly. This is in line with earlier findings on the time course of plausibility effects within sentences (McRae and Matsuki, 2009; Matsuki et al. 2012). Since items were presented phrase-by-phrase, rather than word-by-word, we cannot determine the actual starting point of the effect. By the AdvP, both Low Backshift and High Progression pull apart from High Backshift. This may be a spillover effect of the plausibility relationship established at the verb. However, it may also be an effect of the adverb itself. With a methodology with better temporal resolution, such as eye-tracking while reading, it might be possible to discriminate between effects at the verb and the adverb with the current items. However, if further self-paced reading studies are run, a spillover phrase should be inserted between the V+1 region and the AdvP region to better isolate the effects of the verb.

It is puzzling that the difference between High and Low only emerged between Backshift conditions. It should also have emerged between High and Low Progressions, since at the V+1 region the High conditions were still identical one another, and the Low conditions were identical to one another. However, it should be noted that these effect sizes were not very large, and some—if not all—of them would not survive correction for multiple comparisons.

At the AdvP, there was no difference between Low Backshifts and Low Progressions. This is a different result from the finding for the identical items in Experiment 1, i.e., the Final conditions. In Experiment 1, Final Progressions were slower than Final Backshifts in this region (t=2.26), which was a puzzling result. The difference in Experiment 1 could have been spurious, since the effect size was small and likely not robust under correction for multiple comparisons.

The effects that did emerge at AdvP—both Low Backshifts and High Progressions were read more slowly than High Backshifts—could be interpreted as a reanalysis penalty. In the High Progression cases, comprehenders may have committed to a backshift analysis by the end of the V+1 region, only to have that analysis refuted by the AdvP.

At Spillover1, the results of this experiment once again diverged from those of Experiment 1. In this experiment, there were no differences between Low Backshifts and Low Progressions (nor between any conditions). In Experiment 1, Final Backshifts were slower than Final Progressions.

As for the interaction at Spillover2, it could be a delayed effect of Plausibility and the full temporal relation disambiguation at AdvP. However, in earlier stages of analysis, in which the exclusion criteria were less stringent, this effect did not reach significance. The exclusion criteria issue will be discussed further in the general discussion. At any rate, we will not put too much stock in this interaction.

Overall, the results of Experiment 2 further suggest that backshift is not inherently more costly to comprehend than progression. At the very least, they suggest that any baseline difference between the two is modulated by the various pragmatic cues contained in Plausibility. In fact, the finding that High Progressions were slower than High Backshifts at AdvP suggests that pragmatic cues can lead to a reversal of the relative comprehension costs of backshift and progression. It seems that when readers encounter a verb in S2 that is likely to be in a backshift relationship with the verb in S1, they can leverage this information relatively quickly, as indicated by the emergence of a Plausibility effect at the V+1 region.

2.7 Experiment 1C: Aspect x Temporal Relation

2.7.1 Design and Predictions

Experiment 1C investigates the effects of grammatical aspect on the processing of temporal relations. It had a partially crossed 2x2 design with the factors Aspect {Simple, Perfect} and Temporal Relation {Backshift, Progression}. The Perfect Progression condition was excluded because it is ungrammatical, as illustrated in (23), leaving three conditions. All of the disambiguating AdvPs were Final, and the plausibility of backshift was Low throughout.

(23) Lianne poked Eric. *He **had** pinched her on the arm a moment **later**, in the aisle near the back of the lecture hall.

Since the perfect aspect marker ‘had’ is an unambiguous cue to a backshift relation in the contexts used in the experiment, we would expect it to facilitate backshift processing. Given the results of the plausibility experiment, in which we found a defeasible cue to have effects on processing, we expect an indefeasible cue like perfect ‘had’ to have a similar effect.

2.7.2 Materials

Thirty-six experimental stimuli were combined with the same 72 fillers used in the previous experiments, with comprehension questions constructed and presented as they were in the previous experiments. One item was excluded from the analysis due to typographical error, for a total of 35 items in the final analysis. A sample item is in (24); the complete materials are in Appendix B.3.

(24) **S1:** Lianne poked Eric. |

a. **S2** (Simple, Back): He pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

b. **S2** (Simple, Prog): He pinched | her on the arm | a moment later, | in his seat | near the back | of the lecture hall.

c. **S2** (Perfect, Back): He had pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

2.7.3 Participants

Participant recruitment, payment, and experiment duration were the same as for Experiments 1A and 1B (§2.5.3). Eight participants were excluded from the analysis on the basis of comprehension question performance, and four were excluded for apparent spacebar-holding. The remaining 48 participants are included in the analysis.

2.7.4 Procedure

The procedure was the same as in Experiments 1A and 1B (§2.5.4).

2.7.5 Results

The data were analyzed via the same procedures used for Experiments 1A and 1B (§2.5.5).

The mean residualized RTs for each condition are presented in Figure 2.4. Pairwise comparisons showed that, at the SV region, Simple Backshifts were read more slowly than Perfect Backshifts (β=-.06 (.02), t=-3.3 with reference level Simple Backshift), as were Progressions (β=-.07 (.02), t=-3.91 with reference level Progression). At V+1, the relation between Simple Backshifts and Perfect Backshifts reversed—Perfect Backshifts were read more slowly (β=.05 (.02), t=2.29 with reference level Simple Backshift). At Spillover1, Perfect Backshifts were read more slowly than Progressions (β=.03 (.01), t=2.04 with reference level Progression). The results of the LME analyses are summarized in Tables 2.12-10.

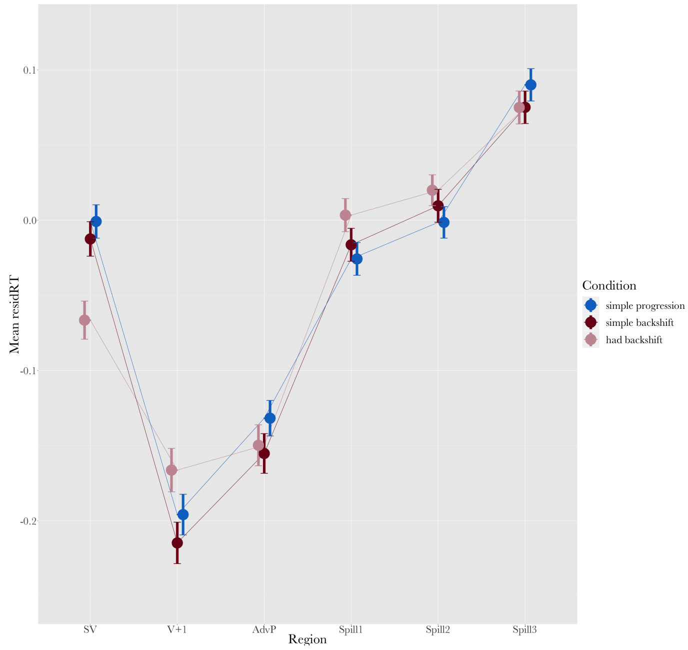


Figure 2.4: Mean residual reading times by region in Experiment 1C. Error bars show standard errors of means.

Table 2.12: Summary of linear mixed effects models for pairwise comparisons in Experiment 1C. Reference level Progression. Bold cells indicate significant coefficients.

	SubVerb*		V+1		AdvP	
	β	t	β	t	β	t
Perf Back	**-.07 (.02)**	**-3.91**	.03 (.02)	1.39	-.02 (.02)	-1.02
Sim Back	-.01 (.02)	-.76	-.02 (.02)	-.92	-.02 (.02)	-1.37
	Spill.1		Spill.2**		Spill.3	
	β	t	β	t	β	t
Perf Back	**.03 (.01)**	**2.04**	.02 (.02)	1.33	-.01 (.02)	-.83
Sim Back	.01 (.01)	.83	.01 (.02)	.61	-.01 (.02)	-.78
*Random slopes for items, full subjects						
**Decorrelated random effects for subjects, full items						

Table 2.13: Summary of linear mixed effects models for pairwise comparisons in Experiment 1C. Reference level Simple Backshift. Bold cells indicate significant coefficients.

	SubVerb*		V+1**		AdvP	
	β	t	β	t	β	t
Perf Back	**-.06 (.02)**	**-3.3**	**.05 (.02)**	**2.29**	.004 (.02)	.2
Prog	.01 (.02)	.76	.02 (.02)	.97	.02 (.02)	1.37
	Spill.1		Spill.2*		Spill.3	
	β	t	β	t	β	t
Perf Back	.02 (.01)	1.2	.01 (.02)	.72	-.001 (.02)	-.04
Prog	-.01 (.01)	-.83	-.01 (.02)	-.61	.01 (.02)	.78

*Random slopes for items, full subjects
**Random intercepts for items, full subjects

2.7.6 Discussion

The effect of the perfect marker *had* at the SV region may have been due to the greater length of the Perfect Backshift condition relative to the other two conditions. The residualization may have overcorrected for this systematic length difference. At V+1, the Perfect Backshifts were significantly slower than Simple Backshifts, but not Progressions. Since the Simple Backshifts and Progressions were still string identical at this region, I am not sure how much should be made of the difference between Backshift conditions, especially since the effect is fairly small.

At Spillover1, the finding that Perfect Backshifts were slower than Progressions, while Simple Backshifts were not, may suggest a delayed cost of processing an overt grammatical cue for a backshift interpretation. It may also reflect an additive cost of both the overt aspectual cue and the backshift-signaling adverb. While it is also conceivable that the adverbial alone was responsible for this effect, we might have also expected a difference between Simple Backshifts and Progressions at this region.

2.8 Interim Discussion

Experiments 1A-1C examined the effects of three factors—adverbial position, plausibility, and aspect marking—on the on-line comprehension of intersentential temporal relations. The results suggest that each of these factors has some effect on on-line temporal interpretation. Of the three, the semantic and pragmatic cues involved in the Plausibility manipulation seem to have the most influence on temporal relation comprehension.

With Experiment 1A, which manipulated the position of a disambiguating temporal adverbial, we attempted to replicate the iconicity effect that Dickey found, i.e., that backshift was more costly than progression, regardless of adverbial position. This experiment did not replicate Dickey's results—there was no significant difference, nor even a trend towards one, between Backshifts and Progressions with Initial AdvPs[12]. For Final AdvPs, there was only a trend towards Backshifts being read more slowly than Progressions in the Spillover1 region (immediately following the AdvP). The failure to replicate may be due in part to differences in the materials. Most prominently, the items in the present experiments were two sentences long, while Dickey's were three sentences long. If the progression interpretation is a default in comprehension, it

[12]Dickey also did not correct for multiple comparisons, so it is possible that the significant differences he found are not robust.

may not be strong enough to show up—at least in self-paced reading—with only one sentence leading into the critical sentence.

Experiment 1A did not contain any highly valid cues for a backshift interpretation. Experiments 1B and 1C introduced such cues, and attempted to get measures of how strongly each modulates temporal relation comprehension, as well as a rough estimate of when the information from those cues seems to be integrated. Experiment 1B manipulated Plausibility of a backshift interpretation. Plausibility is itself an umbrella for several semantic and pragmatic factors that can modulate on-line comprehension (e.g., implicit causality (Rohde, Levy, & Kehler, 2011)). Effects of Plausibility emerged as early as V+1, suggesting that at least some of the factors that contribute to Plausibility can affect comprehension relatively quickly. Obviously, self-paced reading, especially with phrase-by-phrase presentation, cannot tell us exactly how soon the Plausibility effect obtains. Eye-tracking while reading would be one more appropriate methodology for zeroing in on this effect.

Within the Experiment 1B items, the component factors of Plausibility were not balanced—for example, implicit causality verbs were not balanced with non-implicit causality verbs, and causal relations in general were not balanced with non-causal ones. It will take further work, to pick apart the factors that contribute to Plausibility and compare the strength and timing effects to one another. The experiments I present in Chapter 3 approach ‘plausibility’ from the perspective of SDRT, and in the context of discourse representation in memory, rather than on-line comprehension.

Nevertheless, Experiment 1B demonstrated that semantic and pragmatic factors can modulate temporal relation comprehension relatively quickly (for self-paced reading). For various formal accounts of discourse structure that depend on grammatical tense and aspect markers to establish temporal relations, this poses a problem. These accounts do not predict that backshift interpretations can arise without overt grammatical marking; some predict them to be impossible altogether.

Experiment 1C held Plausibility constant (Low), and introduced the perfect aspect marker *had*. In the item context, *had* was an unambiguous grammatical cue to a backshift interpretation. Since this is a highly valid cue in this context[13], one might expect it to level a weak baseline progression advantage. However, this experiment actually found that Perfect Backshifts were read more slowly than Simple Backshifts at V+1, and more slowly than Progressions at Spillover1. It may be the case that there is an added cost of processing an overt grammatical cue. Since Simple Backshifts and Progressions did not differ anywhere—and since Plausibility was always Low, just as in Experiment 1A and Dickey's experiments—this experiment did not provide support for the view that backshifts are generally more costly to comprehend than progressions. At least, they are not more costly in a way that self-paced reading is sensitive to.

[13] John Singler (p.c.) has since informed me that, in African American English, *had* can function as the simple past tense marker, and that feature may be spreading to other American Englishes. I did not collect information about participants' linguistic backgrounds that bears on this, so for now I simply note this as a factor to be monitored going forward.

Based on the findings of all three experiments, we seem to be at or near the limit of what self-paced reading measures can tell us about the effects of these factors on temporal relation comprehension. Eye-tracking while reading would give us a more precise and a more accurate picture of when these factors begin to affect comprehension, and of how long their effects persist. Because of the ongoing COVID-19 pandemic, it has not been possible to run eye-tracking while reading studies. Fortunately, there is another methodology—the Maze task (Forster et al. 2009)—that can be implemented remotely and provide a better picture of incremental integration than self-paced reading.

2.9 Experiment 1D: Plausibility in the Maze

2.9.1 Enter the Maze

The Maze task (Forster, Guerrera, & Elliot, 2009) is intended to force comprehenders into a fully incremental mode of parsing. In this task, comprehenders advance through a sentence word-by-word by making a series of binary choices: they are presented with two words, only one of which—the **target**—is a grammatical continuation of the sentence. The other word—the foil—is the same length as the target.[14] A Maze trial continues as long as the comprehender correctly selects the target words. If they select a foil, the trial terminates. A schema of a Maze trial is shown in 2.14. A compre-

[14]There are variations on the nature of the foil—it may be a nonce word, or simply a string of 'x's of the same length as the target.

hender would see the words in bold, selecting the left-side option by pressing the ‘e’ key and the right-side option by pressing the ‘i’ key. The relative ordering of the target and foil is randomized for each screen.

Table 2.14: Maze schema.

Screen 1		Screen 2		Screen 3		Screen 4	
She	**xxx**	**fuller**	**forgot**	**bit**	**her**	**glasses.**	**allows.**
(target)	(foil)	(foil)	(target)	(foil)	(target)	(target)	(foil)

The idea is that, in order to correctly identify the target, the comprehender must have fully parsed everything they have already seen, and must commit to a parse for the word they select. In principle, the comprehender cannot successfully/reliably perform the Maze task if they delay parsing decisions. In this respect, the Maze is unlike the classic self-paced reading task. The Maze, then, may provide a more accurate and precise way to track integration costs over the course of a sentence than SPR.

The Maze task has been shown to be an effective method for studying a variety of syntactic processing phenomena (Freedman & Forster, 1985; Nicol, Forster, & Veres, 1997; Forster et al., 2009); it is still largely unproven as a tool for studying discourse processing.

2.9.2 Design and Predictions

In this experiment, I implement the Plausibility x Temporal Relation design employed in Experiment 1B in the Maze. First, this represents an attempt to replicate the

SPR results. If this is achieved, we will hopefully also get a clearer idea of exactly when plausibility and temporal relation modulate sentence/discourse comprehension.

The first region of interest is the verb, because it provides the earliest information about the coherence/temporal relation between the eventualities in the second and third sentences. In the SPR task, this critical verb was lumped into a region with the subject. No differences were observed in that task. If there is a plausibility effect this early in the critical sentence, we'll have a better chance of seeing it in the Maze. The next region of interest is the postverbal region, i.e., the region between the verb and the disambiguating temporal adverbial phrase. In the SPR task, there was a significant effect of Plausibility in this region—the Low Plausibility conditions were read more slowly than the High Plausibility conditions. Recall that temporal relation had not yet been explicitly disambiguated at this point—because the temporal adverbials were postverbal, the High plausibility conditions were still identical to one another at this point, as were the Low plausibility conditions.

If backshift always incurs greater comprehension costs than progression, we predict the High plausibility conditions to have slower Maze response times than Low plausibility as early as the verb. This is because comprehenders may be able to infer a coherence relation as soon as they have integrated the verb. That relation may entail either backshift (which is intended in the High plausibility conditions) or progression (the intended inference of the Low plausibility conditions). If comprehenders do not infer a coherence relation until they have integrated the verb and all its arguments, a plausibility effect would not be expected until the postverbal region.

The final region of interest is the temporal adverbial phrase, because it explicitly disambiguates the temporal relation to either backshift or progression. Under the view that backshift is always more costly than progression, we predict that the Maze RTs for backshift adverbs, such as 'earlier', will be higher than for progression adverbs, such as 'later.' If plausibility exerts a stronger influence than temporal relation, we predict that the High plausibility backshifts and the Low plausibility progressions[15] will have lower RTs at the adverb than the other two conditions, in which the plausibility of a backshift interpretation and the explicit temporal relation are incongruent.

2.9.3 Materials

The 64 experimental stimuli were three-sentence discourses. A sample item is in (25). The complete materials are provided in Appendix B.4. Thirty-six of them are expanded versions of the stimuli from Experiment 1A (Plausibility x Temporal Relation). In these stimuli, one sentence has been added at the beginning of the discourse, which establishes that the narrative is set in the past and introduces the character(s). This establishing sentence makes these stimuli more similar to those that Dickey used

[15]Reminder: this means a progression that has low plausibility as a backshift, which makes it a highly plausible progression.

than the original stimuli from Experiments 1A-C.[16]

(25) a. High plausibility Backshift (more coherent): Lianne and Eric sat down next to each other. She poked him. He fell asleep in biology class a moment earlier, in his seat near the back of the lecture hall.

b. High plausibility Progression (less coherent): Lianne and Eric sat down next to each other. She poked him. He fell asleep in biology class a moment later, in his seat near the back of the lecture hall.

c. Low plausibility Backshift (less coherent): Lianne and Eric sat down next to each other. She poked him. He pinched her on the arm a moment earlier, in his seat near the back of the lecture hall

d. Low plausibility Backshift (more coherent): Lianne and Eric sat down next to each other. She poked him. He pinched her on the arm a moment later, in his seat near the back of the lecture hall.

Each item had an accompanying polar comprehension question, which was only displayed for trials in which participants read/Mazed through the entirety of the last sentence of the discourse. The fillers were 48 three-sentence discourses that belonged to another experiment.

[16] I am very grateful to my RA, Spencer Gilbert, who was an enormous help in constructing these stimuli. I am also grateful for the support of the UCSC Linguistics Department Honig RA Program, which funded this RA position.

2.9.4 Participants

All participants provided informed consent, and were asked to report their age and basic linguistic background information. A total of 72 native English speakers were recruited via Prolific. Participants were paid $6 for completing the study. Twenty-four participants were excluded from analysis because they completed less than 70% of the Maze trials, or they scored below 70% accuracy on the comprehension questions. The remaining 46 participants are included in the analysis.

2.9.4.1 Procedure

The experiment was presented through Ibex. The first two sentences of each trial were presented in a centered sentence-by-sentence self-paced reading format. Participants pressed the spacebar to advance. Participants read the third and final sentence of each item in the Maze. They selected what they judged to be the correct continuation word by pressing the 'e' key for the word displayed on the left, and the 'i' key for the word displayed on the right. The left-right ordering of the target and foil words was randomized. If participants Mazed through this sentence in its entirety, they were presented with the comprehension question, and responded 'yes' or 'no' by pressing the 'e' or 'i' keys, respectively. If participants did not reach the end of the final sentence, the trial terminated, and the experiment advanced to the next trial.

Order of presentation was randomized for each participant in a Latin square design.

At the beginning of the experimental session, participants were given written instructions for the task. They then completed seven practice trials. Then, they completed five burn-in trials. These were list-invariant, and are not included in the analysis. The experiment typically lasted 35-45 minutes.

2.9.5 Results

In this document, I present the descriptive results of the Maze task. The first region of interest is the verb. The mean response times for this region, split by condition, are shown in Figure 2.5. Recall that, at this region, the High plausibility conditions are still identical to one another, as are the Low conditions. There is a very slight trend towards the High plausibility verbs having lower RTs than the Low plausibility verbs.

In the graph in Figure 2.5, all the verbs are grouped by condition, regardless of their length. In Figure 2.6, they are broken down by character length. The small High $<$ Low trend suggested by the previous graph does not appear to survive the split. This tracks with the SPR results, in which no differences were observed in the SV region.

The second region of interest is the postverbal region. The RTs for each word position in this region are plotted in Figure 2.7, split by the length (in words) of the region. In this region, as in the verb region, the High conditions are still identical to one another, as are the Low conditions. There do not appear to be any consistent trends at any position within this region. This contrasts with the SPR results. In the SPR task,

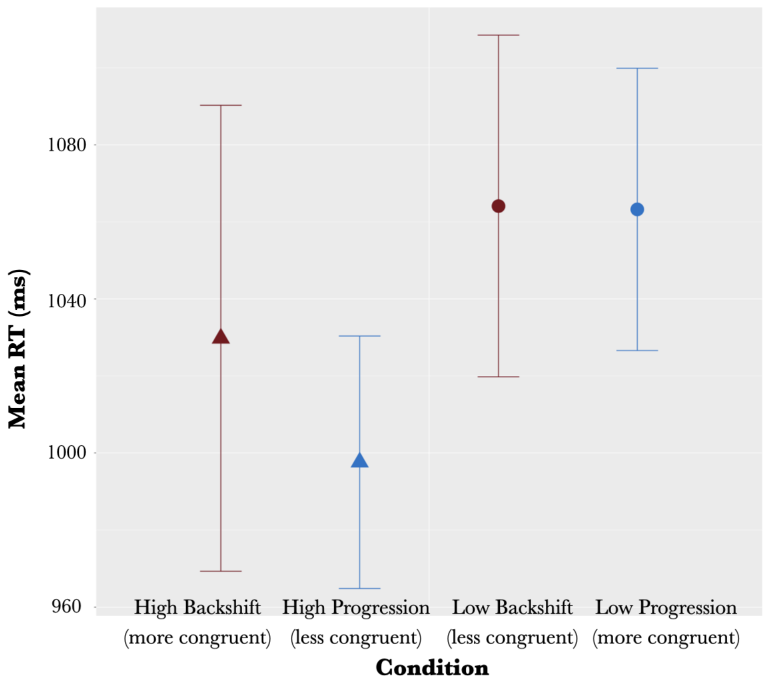


Figure 2.5: Mean response times, Verb region in Experiment 1D. Error bars show a 95% confidence interval.

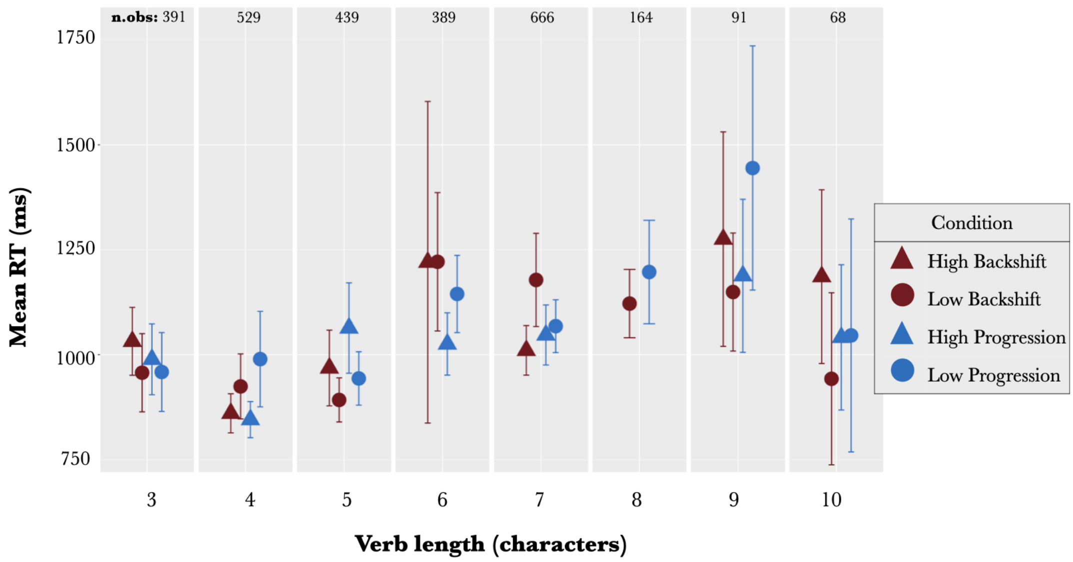


Figure 2.6: Mean response times, Verb region, split by verb length, in Experiment 1D. Error bars show a 95% confidence interval.

this region was presented as a phrasal chunk, and the Low conditions were found to have been read more slowly than the High conditions. A clear potential problem here is that the nature of the postverbal region varied, and wasn't controlled. Some verbs were obligatorily transitive (35 items), some were optionally transitive (26 items, 3 of which were intransitive in the items), and some were intransitive (3 items). Of the transitive items, there was a mixture of full DP, pronominal, and PP arguments.

The final region of interest is the temporal adverbial phrase. Figure 2.8 shows the mean RTs for the full adverbial phrase—the RTs for the individual words in the AdvP were aggregated. This is the region in which the High conditions become differentiated from one another, as do the Low conditions. There is a slight trend towards the High Backshift and Low Progression RTs being faster than the Low Backshift and High Progression RTs. This tracks with the SPR results—in that task, the same pattern emerged as weak but significant.

Finally, Figure 2.9 shows the mean RTs at the point of temporal disambiguation, the temporal adverb itself. The data are split by the specific adverbs used in the stimuli; each adjacent pair always appeared together—for stimuli in which Backshift was signaled by 'earlier,' Progression was always signaled by 'later.' Across all four Backshift adverbs, there is a slight trend towards the High condition having faster RTs than the Low condition. There don't appear to be any consistent trends across the Progression adverbs, or in any comparisons between the Backshifts and Progressions.

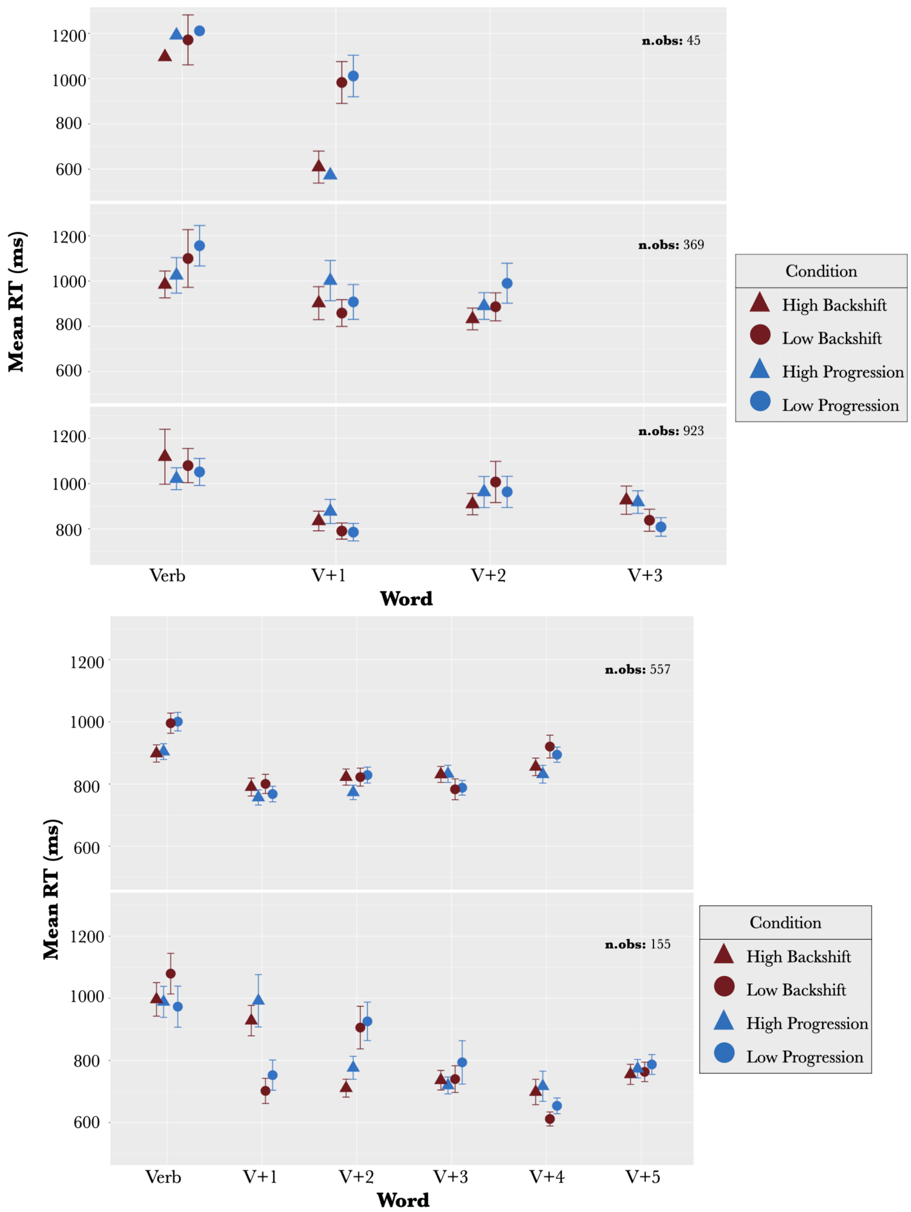


Figure 2.7: Mean response times, Verb Phrase, split by number of words in postverbal region, in Experiment 1D. Error bars show a 95% confidence interval.

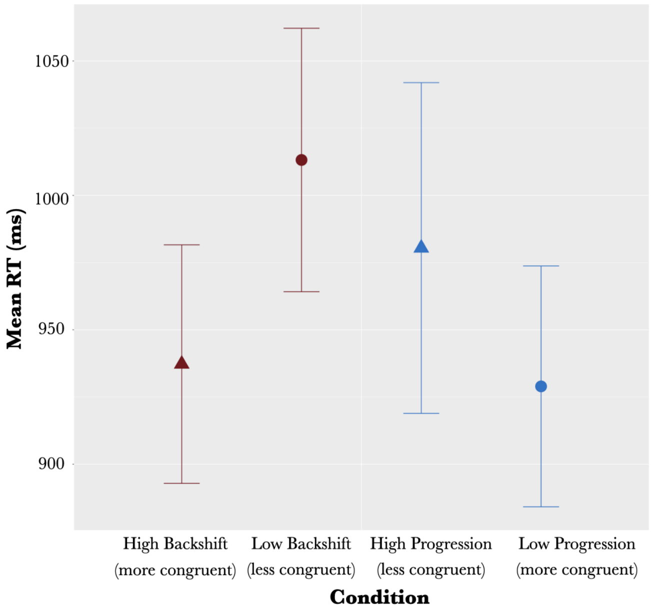


Figure 2.8: Mean response times, Adverbial Phrase, in Experiment 1D. Error bars show a 95% confidence interval.

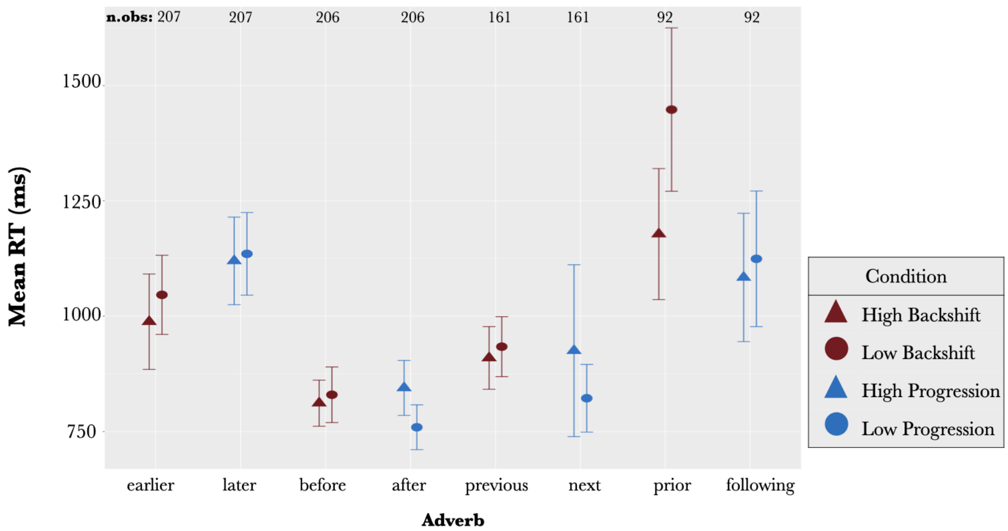


Figure 2.9: Mean response times, Adverb region, split by specific adverbs used, in Experiment 1D. Error bars show a 95% confidence interval.

2.9.6 Discussion

The Maze version of the Plausibility x Temporal Relation study appears not to have replicated the SPR results. It may very well be that the SPR findings were spurious—as previously noted, the effects were small, and thus unlikely to survive statistical correction for multiple comparisons. Besides that, the Plausibility manipulation itself was heterogeneous—a variety of factors that seem to affect plausibility were mixed together in the stimuli, and not balanced.

The apparent lack of iconicity effect in this experiment is perhaps a stronger challenge to Dickey's original finding than Experiment 1B because, as Dickey did, this experiment used three-sentence stimuli. This was meant to give the comprehender a stronger sense that they were reading narrative discourses, and thus to create a better chance for an iconicity/progression advantage effect to emerge. Since it did not, we have another point against the view of progression as a strong default in discourse comprehension.

2.10 Summary

In this chapter, I introduced some previous formal and experimental work on temporal interpretation in discourse. Much of this work is permeated by the idea that temporal iconicity, i.e., progression, is a default in narrative discourse. In fact, many formal theories in this domain explicitly encode this by making unmarked backshifts—that is, unmarked anti-iconic temporal relations—impossible. As observed by coherence theorists such as Asher and Lascarides (2003), this blanket ban means that such theories undergenerate. However, even for systems like theirs that are more permissive of anti-iconicity, the assumption that progression is the default remains. This iconicity assumption also exists in psychological discourse theories—represented here by Situation Model Theory—at a much more implicit level than in formal linguistic theories.

While progression does seem to be a default of some sort—in naturally-occurring discourses it seems to be the most common temporal relation—the question of how this might translate to discourse comprehension and representation in memory has not been much studied. In self-paced reading studies, Dickey (2001) found evidence suggesting that progression is a strong default in discourse comprehension. He found an iconicity effect in the form of faster reading times for progressions than backshifts, which he argued to mean that progression is a strong, first-choice default. That is, the comprehender actively and always predicts progression, and must revise if the relation turns out to be something different, such as backshift.

However, in a set of four experiments, I failed to find an iconicity effect in self-paced reading (Experiments 1A-1C) and the Maze task (Experiment 1D). Although my experiments were not attempts to exactly replicate Dickey's, they had more statistical power and investigated a wider range of factors that might influence the comprehension of temporal relations. When it comes to discourse comprehension, progression does not appear to be a strong default. What these experiments do suggest—especially Experiment 1B—is that the plausibility of a certain temporal interpretation may modulate temporal comprehension. In the next chapter, I explore the interactions between temporal relations and plausibility, focusing in particular on causal relations.

Chapter 3

Temporal distance and causality in Situation Model Theory

In this chapter, I investigate the roles of temporal and causal information through the primary lens of Situation Model Theory (SMT), a psychological theory of discourse coherence and processing. In §3.1, I provide an overview of this theory and a selection of significant research within the framework. In §3.2, I give more detail about the temporal and causal dimensions, which are the focus of the experiments in this chapter. In §3.3, I summarize the two studies that form the jumping-off point for my studies, which I present in §3.4. In these studies, I investigate what counts as a discontinuity in the temporal dimension of SMT, building on previous findings that were interpreted as iconicity effects. In my studies, I found no clear evidence of iconicity effects, though

I ultimately cannot rule them out either. The current results do suggest that the causal dimension has a fairly large role in situation model construction.

3.1 Basics of Situation Model Theory

3.1.1 Why not stick to SDRT?

Asher & Lascarides (2002) expressed some hope that SDRT can serve as, or at least be ported into, a processing model for discourse. They make only brief mention of this, leaving it for others to operationalize, to generate predictions, and to test SDRT in this way. Under a fairly straightforward operationalization, different instantiations of the same discourse relation are not expected to have different effects on a comprehender's mental representations of the discourse and the described situation. For example, let's consider the three versions of the discourse in (1). The final relation in each version is Narration. It could be argued that (1-c) does not form a Narration with the preceding sentence because of the length of time the initial adverbial denotes. This will be discussed more later on, but at any rate, (1-a) and (1-b) are uncontroversially Narrations.

(1) Marge went to the farmer's market. She found some good cheddar cheese first. Then she went through the produce section. She got a bag of peaches. She also bought a big bunch of basil just because it was on sale.

a. A moment later, she decided to make pesto with it.

b. Ten minutes later, she decided to make pesto with it.

c. Ten hours later, she decided to make pesto with it.

Under the proposed view of SDRT, there should be no differences between (1-a) and (1-b) in processing or retrieval performance. However, there is evidence from psychology which suggests that different 'story time' distances, such as those in (1-a)-(1-c), do indeed have different effects on comprehenders' mental representations of described situations (Zwaan 1996, a.o.).

This section provides an overview of discourse-processing research in psychology. Since the early 1980s, several theories of discourse processing have grown out of pioneering work by Johnson-Laird (1983), Van Dijk and Kintsch (1983), and others. These theories differ somewhat in terminology and some details, but, by and large, they are not radically different from one another.[1] So, in the current work, we adopt the terminology of one particularly prominent theory, Situation Model Theory, which flourished in the 1990s, though not all the works we discuss are grounded in it. While some researchers in this domain have engaged formal linguistic discourse theories, for the most part psychologists and linguists have proceeded independently from one another.

In psychology, the overarching question of discourse coherence is very similar to that of formal linguistics, but is concerned in particular with a discourse's representa-

[1]To give one notable example, the term 'situation model' as used in the literature I cover here is equivalent to Johnson-Laird's 'mental model of discourse' (1983).

tion in memory. The critical condition for discourse coherence is that it is 'possible to construct a unitary representation that integrates all the information carried in its separate sentences' (Garnham, Oakhill, and Johnson-Laird 1982:32). So, somewhat more narrowly, the questions that SMT researchers have engaged include the following: (i) What are the components of that 'unitary representation'?; (ii) What factors allow (or disallow) comprehenders to build a coherent representation?; and (iii) How does the structure of that representation affect retrieval?

My focus in this literature overview is mainly on the event-indexing model, a prominent theory under the SMT umbrella (Zwaan, Langston, and Graesser 1995). In this iteration of SMT, it is argued that discourses have three layers of representation in memory. First, there is the surface, or verbatim, representation of the linguistic input. Next, there is the 'textbase', or semantic representation. As conceived of by Zwaan and colleagues, the textbase is a network of propositions, linked by argument overlap. Finally, there is the situation model: a representation of the described situation, which is constructed based on our extralinguistic world knowledge in conjunction with the representations of the linguistic input itself, i.e., the first two layers. In SMT, linguistic content is viewed as 'processing instructions' for constructing a situation model (Zwaan and Radvansky 1998:162).[2] Further, it appears that comprehenders always construct full situation models during discourse comprehension, even if they can meet

[2]These processing instructions need not be linguistic–visual depictions of situations can also serve as situation model instructions.

the task demands simply by relying on the surface and textbase layers of representation (Zwaan, Langston, & Graesser 1995).

3.1.2 Dimensions of Situation Models

The various iterations of SMT largely agree that situation models comprise a handful of 'dimensions,' which comprehenders monitor during discourse processing. Five dimensions are commonly proposed: temporal, causal, spatial, entity (characters and objects), and character goals. Experimental evidence suggests that these dimensions are not fully independent of one another, but that at least some of them do have isolable effects on situation model construction. In online comprehension, comprehenders monitor these dimensions for continuity. If the information about these dimensions provided by some linguistic unit, e.g., a sentence, is continuous with the information provided by the previous sentence, the incoming information is integrated with the existing situation model. If there is a discontinuity in one or more dimensions, situation model update may occur. For instance, a temporal discontinuity may be signaled by an adverbial phrase such as 'a month later.' Three of these dimensions—temporal, causal, and spatial—have been studied in rather more depth than the others. The other dimensions appear under somewhat different names across the SMT literature, sometimes corresponding to somewhat different definitions. They are also sometimes split into more dimensions, but here for simplicity's sake we will use just two: entities (characters and inanimate objects) and character goals.

Under the simplest version of SMT, the comprehender treats all the dimensions equally as regards situation model update. If the information about all the dimensions that is conveyed by some linguistic unit is continuous with the information provided by the previous linguistic unit, then the incoming information is integrated with the existing situation model. If there is a discontinuity in one or more of the dimensions, then situation model update occurs. However, it is reasonable to wonder if all the dimensions really are equal in this respect. Much work in SMT in the early to mid-90s sought to answer this question (sometimes as a sort of a side effect of establishing which dimensions comprehenders monitor in the first place), as well as various contingent questions. Some of the questions SMT has pursued are in (2).

(2) a. Does any discontinuity in any dimension trigger situation model update?

b. What is situation model update? Does a change in one dimension cause the full model–all the dimensions–to be updated? Or is the information from the unchanged dimensions maintained, while a localized update occurs for the changed one?

c. If the dimensions aren't monitored equally closely (i.e., they don't all have the same update-triggering potential), how are they ranked? Can the ranking be modulated by, e.g., task demands?

3.1.3 Previous SMT studies

Zwaan, Magliano, and Graesser (1995) (henceforth, ZMG) studied the temporal, causal, and spatial dimensions. Their main goals were to build on previous evidence that, at baseline, the temporal and causal dimensions are monitored more closely than the spatial. On top of that, they tested whether reading instructions affected that monitoring, and thus situation model construction. (Bravely, they also used naturally-occurring discourses in an effort to show that SMT isn't limited to experimenter-constructed texts.) Participants were asked to read two short stories (including one by Gabriel García Márquez) in sentence-by-sentence SPR. After reading both, they were asked to write a summary of each. The only factor, reading instruction, was between-subjects. Half of the participants were instructed to read 'normally,' i.e., as they would for pleasure. The other half–the 'memory' group–were instructed to read so as to be able to give a vivid and detailed account of what happened. ZMG expected the 'normal' condition to serve as a baseline, and that a situation model update triggered by a dimensional discontinuity would be signaled by a reading time penalty in the discontinuous sentence. For the memory condition, they predicted that, if the reading instructions prompted maximal situation model construction (i.e., constant monitoring of every dimension), there should be a reading time penalty for any discontinuity. Alternatively, if the instructions prompted more focus on the surface form of the text than on the situation model, any reading time penalties should be smaller than the corresponding ones in the ‘normal’ condition (if they appear at all).

The investigators coded each sentence in the stories as +/- continuous with the preceding sentence on all three dimensions. (They also coded each sentence for several other variables, including number of syllables and number of new argument nouns, that were partialed out in the analysis.) In the ‘normal’ instructions group, they observed significant discontinuity penalties in reading times for both temporal and causal discontinuities. This suggests that, by default, comprehenders always monitor temporal and causal continuity relatively closely. In contrast, they seem to monitor spatial continuity less closely. However, it appears that dimension-monitoring levels depend at least in part on task instructions/demands–in the ‘memory’ instructions group, no significant discontinuity penalties emerged. In Zwaan et al.’s interpretation, the instruction effect is one of attention allocation–in the ‘memory’ condition, comprehenders allocated more attention to constructing the text’s surface representation than the situation model. Under this view, their study manipulated the amount of attention paid to different layers of the text representation. It does not manipulate attention allocation within situation model construction alone. A number of subsequent studies, among them Therriault, Rinck, and Zwaan (2006), attempt to do exactly that.

In their first experiment, Therriault, Rinck, and Zwaan (2006) (henceforth, TRZ) employed several of the same design elements as ZMG, including naturally-occurring narratives–coded using ZMG’s methods–and a between-subjects instructions manipulation. Each version of the instructions asked participants to attend to one situation model dimension (spatial, temporal, character), and each group was presented with comprehension questions about that dimension only. These groups were compared to a control group which was basically the same as ZMG's 'normal' group. TRZ predicted that temporal and character shifts should always incur RT penalties, regardless of instructions, while spatial shift effects should be (more) susceptible to the instructions manipulation. They found that temporal shifts and character shifts did indeed incur RT penalties across all four instructional groups, while the spatial shifts did not occur in the character instruction group.

TRZ's second experiment used constructed materials and an expanded design. The same between-subjects conditions (spatial, temporal, character instructions and comprehension questions) were used. In addition, there were eight within-subjects conditions: all combinations of shift/no-shift on all three dimensions. Once again, TRZ found temporal and character shift penalties across all instructional groups; penalties were largest in their corresponding instructional groups. The spatial shift penalty, in contrast, only occurred in the spatial instruction group. The penalties for individual dimensions also appeared to be additive–RTs increased as the number of shifted dimensions increased.

Based on the results of ZMG, TRZ, and various others (Komeda and Kusumi, 2006; Ezzyat and Davachi, 2011; Bailey, Kurby, Sargent, and Zacks, 2017; a.o.), it seems that readers monitor the temporal, causal, and character dimensions all the time during discourse comprehension, regardless of task demands. Shifts in these dimensions appear to automatically/always trigger situation model updates. Shifts in other dimensions do not seem to have this effect on situation model construction. (But c.f. Radvansky and Copeland (2010), who hypothesize that spatial shift effects do not show up in RTs because they are actually much easier to process than other shifts.)

3.2 Two important dimensions

3.2.1 The temporal dimension

Like formal coherence theories, SMT identifies temporal information as a significant factor in discourse-structuring and the mental representation of discourse. The temporal dimension is perhaps the most investigated aspect of situation models and their construction, and is also one of the main foci of the current work. There seems to be general consensus in SMT that temporal discontinuities always constitute boundaries in situation models, i.e., they always trigger situation model update. Situation model update has been shown to include all dimensions, even when there is only a discontinuity in one of them. Various probe recognition studies–which I will discuss in more detail below–have found that comprehenders are worse at recognizing information from non-temporal dimensions if there is a temporal boundary between the probe presentation and the original appearance of the probe word (Zwaan, 1996; Rinck and Bower, 2000; a.o.).

One significant question in the temporal domain is how temporal distance affects the accessibility of other information in situation models. To answer it, one must determine what kind of time is relevant–the narrative-internal time, or the narrative-external time. (Narrative-internal time is the temporal dimension of the situation model, while the external time is basically the reading time.)

Additionally, the question of what counts as a temporal boundary is not fully settled. In quite a bit of SMT work, there is a tacit assumption that backshifts are always boundaries. SMT research on the temporal dimension has almost exclusively focused on progression relations, which in experimental stimuli are commonly marked by sentence-initial temporal adverbials, e.g., 'An hour later.' Among such progressions, the findings differ as to what amount of story-time, i.e., the length of time denoted by the adverbial, constitutes a boundary. Some studies, including Rinck & Bower (2000), have found that adverbials that denote relatively short story-times, typically minutes, do not seem to count as boundaries. In contrast, they found that adverbials denoting longer story-times, typically two or more hours, do seem to constitute boundaries. However, their findings do not align with those of Zwaan (1996). In Zwaan's studies, only the shortest possible story-time progression, i.e., 'A moment later,' did not count as a boundary; progressions ranging from minutes to hours and days all behaved like boundaries.

3.2.2 The causal dimension

In SMT, causal links between events are considered a significant factor in situation model construction. Work on the causal dimension has focused mainly on causal progressions, i.e., Results. Not all of this work is couched in the SMT framework, though SMT is influenced by it. In a fairly early study of discourse comprehension, Mandler (1986) found that temporal iconicity violations gave rise to a reading time penalty, but that causal or enabling relations ameliorated that penalty. Several other studies have found that cued recall for previous discourse material is better when there is a causal link between the most recent material and the material to be recalled than when there is not (Myers, Shinjo, and Duffy, 1987; Duffy, Shinjo, and Myers, 1990; a.o.).

There is an additional nuance to these findings: recall did not increase monotonically as the degree of causal relatedness increased. In fact, recall for material following a 'moderately' strong causal link was better than recall following both weak causal links and very strong causal links.[3] This recall pattern has been attributed to differences in

[3]A sample of the different link strengths from Duffy et al. (1990):

a. Strongly related: Joey's brother punched him again and again. The next day his body was covered with bruises.

b. Moderately related: Joey got angry with his brother in a game. The next day his body was covered with bruises.

c. Weakly related: Joey went to play basketball with his brother. The next day his body was covered with bruises.

the probability of making a causal bridging inference. That is, for a low-related pair of sentences, comprehenders are relatively unlikely to form a causal bridging inference, or a series of such inferences. For a highly related pair, comprehenders are somewhat more likely to form a causal bridging inference than for a low-related pair, but still less likely to do so than for a moderately related pair; in these cases, the inference may not be necessary to forming a causal link between the sentences. For the moderately related pairs, causal bridging inferences are necessary to integrate the two sentences, but aren't too numerous or implausible to draw.

In what may be a side effect of SMT's assumptions about temporal boundaries, Explanations have received relatively little attention. That is, backshifts are generally considered temporal boundaries by default in SMT, and SMT researchers are perhaps not as attuned to the Result/Explanation pair as coherence theorists in linguistics. There has been some work on the comprehension of Explanations marked by 'because' (Millis and Just, 1994; Traxler, Bybee, and Pickering, 1997). However, these studies are concerned with online comprehension and the distinction between incremental and delayed integration. As a result, I do not know of any explicit (published) predictions about the effects of the Explanation relation on situation model construction. However, given the existing SMT and SMT-adjacent work on causal links, combined with SMT's findings and predictions about the temporal dimension, we can make some reasonable extrapolations.

We might expect that Explanations constitute a boundary in situation model construction because they involve a temporal discontinuity. However, any boundary penalties may be mitigated by the causal link of Explanation. So, if we find that Explanations behave comparably to Results qua situation model boundaries, we will have more evidence that the effects of temporal discontinuities can be ameliorated by causal continuity. This would also lead us to predict that backshifts without causal continuity should count as boundaries, at least in comparison to Explanations. (This prediction is not tested in the current chapter.) If Explanations seem to count as boundaries while Results do not, we will have evidence that the temporal discontinuity of backshift relations cannot be fully mitigated by causal continuity. Finally, SMT does not seem to make any clear predictions about how story-time distance might interact with backshifts (with or without causal links). It may be that a backshift and a story-time discontinuity have additive effects, such that a backshift signaled by a short story-time (e.g., 'a moment earlier') creates smaller boundary penalties than one signaled by a longer story-time (e.g., 'a day earlier'). However, SMT does not straightforwardly predict this, though it also does not predict that story-time distance and backshift should not interact.

So, while psychology has gained substantial insight into online discourse comprehension, there is still a comfortingly vast amount to do in this realm. Besides more careful controlling of linguistic factors, there are various specific lines of investigation that psychologists may overlook, but that are of interest to coherence theorists (and other linguists). For instance, there has been little investigation of the role of causal inferences in comprehenders' treatment of temporal discontinuities. Also, the processing of backshifts has been largely ignored in the situation model literature, while in the coherence theory literature, backshifts are a key part of the empirical terrain. As a final example, there is still a lot to investigate regarding the ways in which narrative-internal temporal distance modulates the accessibility of situation model information and linguistic information. Coherence theorists may ask, for instance, whether and how narrative distance modulates pronominal anaphora resolution.

3.3 Two important experiments

In §3.3.1 and §3.3.2, I outline two previous SMT studies, whose findings make conflicting suggestions about what counts as a temporal discontinuity. One goal of my studies is to adjudicate between these two positions.

3.3.1 Rinck & Bower (2000)

Rinck and Bower (2000) (R&B) used a map-learning and probe recognition task to investigate the effects of internal vs. external time on memory for non-temporal information in a situation model. Participants first learned a map of named rooms containing labeled objects. Then, in SPR, they read narratives about a protagonist moving around the memorized layout. At the end of the narrative, participants answered a probe question of the form, 'Is object X in room Y?' R&B used a 5-condition (2x2 + 1) design. Internal time was manipulated with temporal adverbial phrases—'ten minutes later' for the 'short' condition and 'two hours later' for the 'long' condition. External time was based on the number of sentences intervening between the probe presentation and the probed sentence—two interveners in the 'short' condition, five for 'long.' The final condition was a control, in which the probe was presented immediately after the probed sentence. A sample item and map are in Figures 3.3.1 and 3.3.1. It is also notable that the probe location and object are not explicitly mentioned anywhere in the text. The probed sentence describes the protagonist moving from one room to another, which in the map are separated by a third room. The probe is about that third room and the object in it.

R&B found no effect of external time on probe response times, but did find an effect of internal time. The probe RTs for 'long' internal time ('two hours later') were significantly higher than for 'short' internal time ('ten minutes later'). R&B's results suggest that readers can use narrative-internal time to retrieve information from a situation model independently of external time. They also suggest that shifts in the temporal dimension cause all dimensions to be updated, as temporal shifts affected the accessibility of information in the spatial and character-object dimensions. Finally (and perhaps most strikingly), R&B found this temporal shift effect for inferred locations/objects, not simply locations/objects that are explicitly mentioned.

Calvin was one of the janitors at the research center.

Tonight he slowly changed into his work clothes in the wash room.

He didn't like the job much, but he had to keep it because he needed the money to stay in architecture school.

When he opened his locker, he noticed a note: Director of center has misplaced top secret report, must be found immediately!

He would have to make a thorough search of the center during his shift.

He went into the repair shop, but he couldn't see any papers there.

So he walked from the repair shop into the experiment room.

First Intervening Sentence

This room was a big mess, and Calvin would have to clean it up before he could go on.

Second to Fourth Intervening Sentence

Looking around, Calvin thought that someone must have had a party in here.

He saw empty pizza boxes, Coke cans, bottles of beer, and bits of pop corn everywhere.

There was also a puddle of beer on the floor because someone had dropped a bottle.

Final Intervening Sentence

Minutes: After ten minutes, Calvin was finally done cleaning up the room.

Hours: After two hours, Calvin was finally done cleaning up the room.

Test Probe

BED–LOUNGE

Figure 3.1: A sample item from Rinck & Bower (2000).

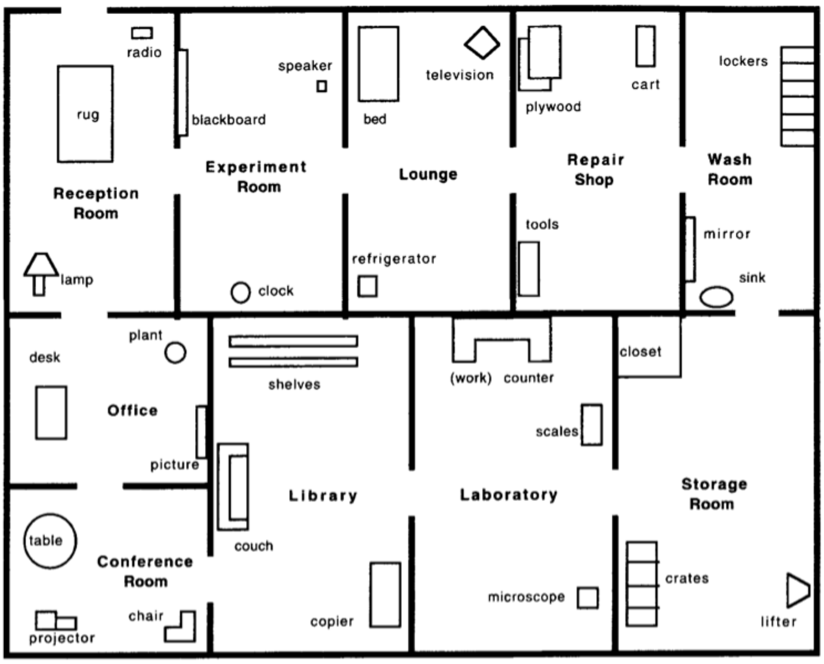


Figure 3.2: A sample map from Rinck & Bower (2000).

R&B's study also address a subpart of the question of how situation models are structured/updated–for each dimension, what counts as a discontinuity/boundary? In regards to the temporal dimension, one hypothesis is that any jump forwards or backwards in time constitutes a discontinuity. After all, humans only experience time unidirectionally and incrementally (that is, moment-by-moment). Under this view, one would expect that 'a moment later' does not trigger situation model update and the resultant decrease in the accessibility of information before the boundary, while anything larger than 'a moment' would. This hypothesis is not tested in R&B's design, since they did not test 'a moment later' at all. However, under this view, R&B's temporal adverbials would both be expected to signal situation model boundaries, and would not necessarily be expected to behave differently in that respect. Since R&B did find that 'two hours later' seemed to constitute a boundary, while 'ten minutes later' did not (or at least a relatively more permeable boundary), their results to at least suggest that the aforementioned hypothesis cannot provide the full picture. Their findings suggest that, in addition to—or perhaps independently from—strict incrementality (i.e., moment-by-moment-ness), temporal shifts may be identified by scenario boundaries. These scenario boundaries would be determined by the comprehender's intuitions about how long the described scenario typically lasts. The comprehender may judge that cleaning a room, to draw from R&B's sample item, usually takes longer than ten minutes, but less than two hours. A forward leap of ten minutes, then, would still be within the boundary of the room-cleaning scenario, while a leap of two hours would cross that boundary, thus triggering a situation model update.

3.3.2 Zwaan (1996)

The comparison that R&B did not make, i.e., that of the boundary-inducing potentials of ‘a moment later’ and any larger forward leap in time, was made by Zwaan (1996).

The experiments in this chapter follow up on Zwaan’s. In his experiments, Zwaan engaged the question of how story-time distance between two consecutive events in a narrative affect their online comprehension and their situation model representations. He also sought to precisify what ‘temporal continuity’ means in SMT. To do this, he compared two views of temporal continuity. The first, the Strong Iconicity view, builds explicitly on Dowty’s (1986) Temporal Discourse Interpretation Principle. The competing view, the Scenario Model, posits that comprehenders determine temporal continuity based on their knowledge of the typical duration of the described scenario (Anderson, Garrod, and Sanford 1983). Under Strong Iconicity, any temporal skip should constitute a boundary, whereas under the Scenario Model, only temporal skips that move the narrative beyond what the comprehender thinks is the typical duration of the described event should count as boundaries. For example, if the relevant event is ‘watching a movie,’ a temporal skip like the one denoted by ‘ten minutes later’ should not be a boundary, while ‘seven hours later’ should be (Zwaan ibid.:198).

Zwaan tested these views via a self-paced reading task with probe recognition. His experiment had one factor, Story-time Distance, with three levels: short-distance, always signaled by ‘a moment later’; medium-distance, always a shift within the typical scenario boundary (e.g., ‘an hour later’); and long-distance, always a shift beyond the typical scenario boundary (e.g., ‘a day later’). Participants read 13-sentence texts sentence-by-sentence; the critical 9th sentence contained an initial temporal adverbial signaling the story-time distance. Immediately after that sentence, participants responded to a probe recognition question–they had to say whether or not they had seen a given word. For positive probes, that word was always drawn from the sentence immediately preceding the critical sentence. Zwaan found that the probe response times were faster following the short-distance adverbials than the medium- and long-distance adverbials, which did not differ from one another. Though the difference was significant, it was quite small (65ms in raw RTs). He interpreted his findings as consistent with Strong Iconicity, and inconsistent with the Scenario Model.

The results of Zwaan (1996) and Rinck & Bower (2000) are in an interesting tension with one another. R&B found a difference between short and long temporal leaps, while Zwaan did not. Perhaps, though, the R&B studies and Zwaan studies cannot fairly be compared. The tasks were quite different, for one thing–R&B’s involved map memorization, and was a probe verification task, not probe recognition like Zwaan’s.

What's more, some of the crucial regions of the items themselves seem to differ. Unfortunately, we can't tell if these were differences between all of R&B's items and all of Zwaan's items, since neither of the papers includes the full item sets. In R&B's items, the boundary sentence has a stative predicate ('After ten minutes, Calvin was finally done...'); Zwaan's boundary sentences have eventive predicates ('An hour later, the telephone rang'). Uncontrolled linguistic factors like this may also have contributed to the differences between R&B's results and Zwaan's.

Given the independently established importance of the temporal dimension in discourse comprehension and memory representation, it seems desirable to have a more detailed understanding of it. Thus, we have reached the jumping-off point for my experiments: the previous studies are not perfectly comparable with one another because they had different designs–including different numbers of levels in the story-time distance factor. Also, they follow the familiar SMT pattern of only testing Narrations.

3.4 Experiment 2

3.4.1 Design and Predictions

3.4.1.1 Strong Iconicity vs. the Scenario Model

These experiments had two main purposes: to adjudicate between the Strong Iconicity view and the Scenario Model, and to expand on the studies outlined above by including backshifts and causal coherence relations. Following Zwaan (1996), the story-time distance factor in the current experiments had three levels—short, medium, and long—so that the hypotheses being compared could really be differentiated. Experiment 2A is an attempt to replicate the Zwaan (1996) results; it focuses only on Narration. Because we also wanted to test backshifts, we needed two coherence relations that form a near-minimal pair on the temporal dimension. Result and Explanation form a better near-minimal pair than Narration does with any of the backshift-entailing relations, so Experiments 2B and 2C focus on Results and Explanations, respectively. Following Zwaan and various other SMT studies, the current experiments employed sentence-by-sentence self-paced reading with a probe recognition task and a comprehension question for each experimental discourse. The dependent measures were probe question response time and probe question response accuracy.

The Strong Iconicity view, which Zwaan's results supported, predicts that any temporal move besides a progression of only one moment should constitute a situation model boundary. For Narrations (Exp. 2A), this view predicts that probe question response times should be split such that the short-distance conditions are answered more quickly than the medium- and long-distance conditions (short RT $<$ medium RT = long RT). The strong iconicity view further predicts that probe question response accuracy should be split such that the accuracy for short-distance conditions is higher than for the other conditions (short Accuracy $>$ medium Accuracy = long Accuracy). We expect the same for Result (2B) as for Narration.

For Explanations (2C) in isolation, this view does not make any clear predictions, because Explanation entails a strong iconicity violation. This view at its simplest does not predict different penalty sizes, though it could be supplemented such that penalty size increases with the size of the temporal leap. However, Strong Iconicity does predict that all the Explanations should have higher probe question RTs and lower probe question accuracy than the short-distance Results and Narrations, simply because Explanations always violate strong iconicity. No further distinctions between Explanations and Results/Narrations are straightforwardly predicted here.

The Scenario Model predicts, for both Narrations (2A) and Results (2B), that both the probe response times and probe accuracy rates for the short- and medium-distance conditions should pattern together, to the exclusion of the long-distance condition. This is because only the long-distance condition should constitute a shift beyond the expected runtime of the described event. Thus, under the Scenario Model, we predict short RT = medium RT $<$ long RT for the probe response times, and short Accuracy = medium Accuracy $>$ long Accuracy for the probe question accuracy.

The Scenario Model, like the Strong Iconicity view, does not make clear predictions about Explanation (2C), but this seems to be because its developers—all psychologists—did not give much explicit consideration to backshifts. If we interpret the Scenario Model as penalizing all Explanations, we predict the probe RT and probe accuracy for Explanations to be worse than those for short- and medium-distance Results and Narrations.

Additionally, we might expect any story-time distance effects we observe to be fairly small. The effects in Zwaan (1996) and Rinck & Bower (2000) were not especially large, and those were for Narrations. If it is true that causal continuity can ameliorate the effects of temporal discontinuity, any differences we see for Results and Explanations could very well be tiny, if they appear at all.

3.4.1.2 SDRT predictions

Finally, let us return to potential operationalizations of SDRT, and the predictions that we might glean from them. Under a view in which all instances of a coherence relation are proposed to have the same effects on discourse processing, SDRT would not straightforwardly predict any effects of story-time distance on situation model structuring. If a sentence is in the same coherence relation with preceding sentence regardless of the story-time distance between them, then those distances are not expected to have different effects on the accessibility of preceding discourse information. However, if a story-time distance is so long that a comprehender actually does not interpret the critical sentences as being in any coherence relation, we may expect to see higher probe RTs and lower probe question accuracy for that distance. When it comes to potential differences between coherence relations, two possibilities readily emerge. One is that ‘accessibility’ in SDRT, i.e., accessibility in the discourse structures one builds in SDRT, corresponds pretty directly to accessibility for retrieval. If this is the case, SDRT predicts no differences in probe task performance between Narration, Result, and Explanation for this design. This is because the probe words are in the sentence immediately adjacent to the sentence containing the story-time distance manipulation. Alternatively, if ‘accessibility’ in SDRT is more about the semantic/pragmatic consequences of establishing coherence relation, rather than the theory-internal structural consequences, the predictions do not really depend on SDRT, but coincide with the predictions we see in the SMT context.

3.4.2 Materials

3.4.2.1 Discourses

Sample materials for Experiments 2A, 2B, and 2C are shown in (3), (4), and (5), respectively. The complete materials are provided in Appendix C. Experimental materials comprised 30 item sets. The stimuli were short narrative discourses, ranging from 6-11 sentences in length. In what follows, I will refer to the subparts of the discourses as labeled in (3).

(3) Experiment 2A: Narration

a. **Lead-in**: Jodie spent the day fishing on Saturday.

She went to her favorite spot at the lake behind her house.

In the morning, she caught a few small perch.

She decided to target bigger fish after lunch.

At around 1 PM, she tied a new hook onto her line.

b. π**1**: Then she put fresh bait on the hook.

c. π**2**: {A moment/Five minutes/A day later}, she cast towards a deeper part of the lake.

d. **Lead-out**: Her strategy worked, and soon she had several large trout in her cooler.

She kept one for her dinner, and shared the rest with her neighbors.

(4) Experiment 2B: Result

Olive had a miserable couple of days last week.

Her acting career had been steadily declining, but last week was especially bad.

First, she learned that the series she was on was going to be cancelled.

Her prospects for new jobs weren't looking very good.

She found out that she had lost a coveted role to a younger actor.

{A moment/A half hour/A day later}, she fired her longtime agent, a respected Hollywood power player.

She began calling up old producer friends herself to see if they had anything for her.

Unfortunately, she was too proud to accept the ‘aging mother’ roles they offered up.

(5) Experiment 2C: Explanation

On Thursday, Lottie had the house to herself overnight.

Her housemates were out of town.

She was excited to have some peace and quiet.

However, she was also a little nervous to be all alone at night.

First, she made sure the doors were locked.

Then, she closed and latched all the windows.

She flipped on the lights on the back patio.

{A moment/Five minutes/Several hours earlier}, she had heard a weird scratching noise in the yard.

The lights revealed a young raccoon.

It was trying to open the trash bin.

The lead-in portion of each discourse ranged from two to seven sentences in length. The lead-in length was varied to discourage participants from discerning which sentence would be followed by the probe question, which could allow them to strategically focus specifically on verbatim recall of that sentence. Additionally, the lead-ins included sentence-initial temporal markers such as ‘then,‘ ‘after that,’ and ‘two hours later,’ so that a sentence-initial temporal marker was not fully predictive of the probe/probe question locations. The experimental manipulations occurred in the $\pi 1$ and $\pi 2$ sentences.

In Experiment 2A, $\pi 1$ and $\pi 2$ were always in a Narration relation. The initial temporal adverbial in $\pi 2$ always ended with 'later.' In Experiment 2B, $\pi 1$ and $\pi 2$ were always in a Result relation. As in Experiment 2A, the initial temporal adverbial in $\pi 2$ always ended with 'later.' In Experiment 2C, $\pi 1$ and $\pi 2$ were always in an Explanation relation. In that experiment, $\pi 2$ the initial temporal adverbial phrase always ended with the word 'earlier.'

The story-time distance factor had three levels—short/medium/long—as exemplified above. The lead-out was always two sentences long, and served to wrap up the narrative and separate the probe question from the comprehension question. Across all three experiments, the lead-ins and $\pi 1$ sentences were held constant as much as possible, so that the results would be maximally comparable to one another. Twenty out of the 30 experimental discourses were used in both Experiments 2B and 2C—they were identical at least through $\pi 2$, except for the direction of the $\pi 2$ temporal adverb ('earlier' vs. 'later'). Lead-outs were also held as constant as possible. Changes were made between experiments only as much as necessary to maintain the overall coherence of each discourse.

3.4.2.2 Probe recognition questions

The probe recognition question was presented after $\pi 2$. Experimental probes were all verbs in the past tense. The probes were evenly split between positive (did appear in the discourse) and negative (did not appear in the discourse). All of the positive probes were the matrix verb of $\pi 1$. Negative probes were matched with the verb of $\pi 1$ for number of syllables and frequency, and were not synonyms of the $\pi 1$ verb.

3.4.2.3 Comprehension questions

The comprehension question was presented after the lead-out. Two-thirds of these questions were designed to be easy. For example, the question for (4) was ‘Does Olive want to play ‘motherly’ types?’. The response accuracy for the easy questions was an exclusion criterion for the data analysis. The hard questions, which included questions about the order of the events in a discourse, were designed to encourage participants to try to comprehend the narratives more ‘deeply,’ rather than to simply scan for potential probe words.

3.4.2.4 Fillers

Twelve filler discourses were constructed to be of similar length and narrative structure to the experimental discourses. The filler probes—both positive and negative—included nouns, adjectives, and verbs. Experiment 2C was run with these fillers, but, due to an experimenter error, both Experiments 2B and 2C ran without them. As such, the results of Experiment 3A may not be as directly comparable with Experiments 2B and 2C as the latter two are with one another. This is discussed further in §3.5.

3.4.3 Participants

All participants provided informed consent, and were asked to report their age and basic linguistic background information.

Experiment 2A (Narrations): A total of 72 native English speakers were recruited via Prolific. Participants received $7 USD for completing the study. 20 participants were excluded from analysis for low accuracy on the easy comprehension questions (<75%). The remaining 52 participants are included in the analysis.

Experiment 2B (Results): A total of 72 native English speakers were recruited via Prolific. Participants received $7.50 USD for completing the study. Seven participants were excluded from analysis for low accuracy on the easy comprehension questions (<75%). The remaining 65 participants are included in the analysis.

Experiment 2C (Explanations): A total of 84 native English speakers were recruited from UCSC undergraduate linguistics courses. Participants received course credit for completing the study. 16 participants were excluded from analysis for low accuracy on the easy comprehension questions (<75%). Two further participants were excluded from analysis because, in the post-experiment debriefing, they correctly identified the experimental manipulation. The remaining 66 participants are included in the analysis.

3.4.4 Procedure

The experiment was presented through Ibex (Drummond 2013). The stimuli were presented in a sentence-by-sentence self-paced reading format. Participants pressed the spacebar to advance; with each press, the next sentence replaced the previous one in the center of the screen. Immediately after advancing past the $\pi 1$ sentence, which contained the critical AdvP, participants were presented with the probe recognition question, 'Did you see the word PROBE?', and were asked to respond 'yes' or 'no' by pressing the 'f' and 'j' keys, respectively. Then, participants read the two-sentence lead-out in the same self-paced reading format. Finally, participants were presented with the comprehension question, and again responded 'yes' or 'no' by pressing the 'f' or 'j' keys. Both questions timed out after being displayed for 5000ms with no response. For both questions, if participants responded incorrectly or timed out, they received the message, 'Wrong. Press the spacebar to continue.' No feedback was given for correct responses. Order of presentation was randomized for each participant in a Latin square design. At the beginning of the experimental session, participants were given written instructions for the task. They then completed three practice trials. The experiment typically lasted 20-35 minutes.

3.4.5 Results

Data were analyzed in the R programming environment (R Core Team 2013), and modeled in Bayesian mixed effects linear regression models using the brms package (Bürkner, 2017; Carpenter et al., 2017). The default priors were used. Models included random slopes and intercepts by participants and items as group-level effects.

3.4.5.1 Reading Times

Mean reading times for all three experiments are plotted in Figure 3.3.

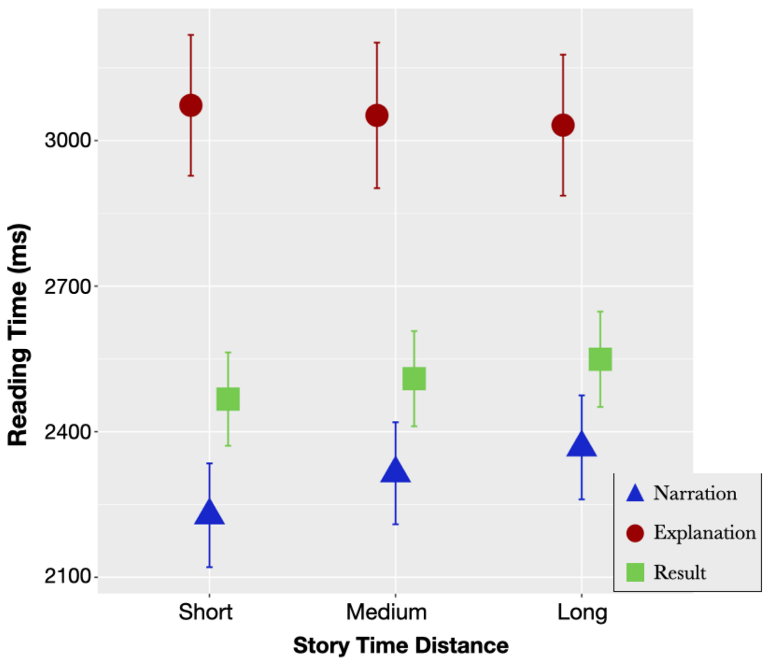


Figure 3.3: Mean reading times, raw, in Experiments 2A-2C. Error bars show a 95% confidence interval.

The data from each experiment were first analyzed separately for story-time distance effects. Reading times shorter than 500 ms or longer than 7000 ms were excluded. The reading times were then corrected for length differences (in number of words). The corrected reading times were entered into the models. Fixed effects estimates are given in Table 3.4.5.1. No story-time distance effects emerged in any of the experiments.

Table 3.1: Story-time distance analysis, reading times: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Short	Intercept	-55.47	56.44	-166.91	54.83
		Medium	80.75	51.99	-19.73	182.02
		Long	79.32	55.04	-28.51	187.90
	Medium	Intercept	22.10	54.95	-85.20	131.25
		Long	-0.14	54.90	-109.43	106.44
Result	Short	Intercept	-11.06	72.17	-152.36	127.68
		Medium	38.49	43.88	-48.36	125.65
		Long	25.72	51.66	-78.85	124.71
	Medium	Intercept	29.43	70.31	-109.38	167.88
		Long	-9.53	52.31	-113.00	93.56
Explanation	Short	Intercept	3.19	78.82	-147.59	159.16
		Medium	56.85	83.79	-101.49	221.23
		Long	-52.99	84.21	-217.45	114.06
	Medium	Intercept	56.65	80.27	-105.80	214.87
		Long	-110.24	88.18	-284.49	58.32

The data from the three experiments were then combined for analysis of coherence relation effects. The data from each experiment were pooled across story-time distances for this part of the analysis. Fixed effects estimates are given in Table 3.4.5.1. Although it appears in the raw reading times (Figure 3.3) as though Explanations were read more slowly than Narrations and Results, no reliable effects emerged.

Table 3.2: Coherence relation analysis, reading times: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Intercept	-18.09	44.40	-111.09	65.53
	Result	1.50	55.82	-110.74	109.35
	Explanation	40.03	63.09	-83.84	163.31
Result	Intercept	-17.10	55.46	-126.07	92.41
	Explanation	32.34	56.65	-81.35	143.17

3.4.5.2 Probe Response Times

Mean response times for correct probe question responses for all three experiments are plotted in Figure 3.4.

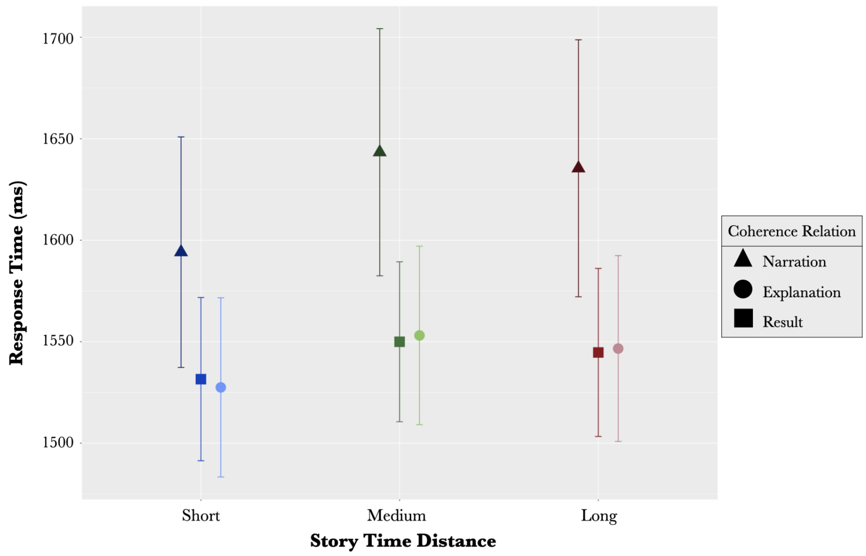


Figure 3.4: Mean probe response times, raw, in Experiments 2A-2C. Error bars show a 95% confidence interval.

The data from each experiment were first analyzed separately for story-time distance effects. Fixed effects estimates are given in Table 3.3. No story-time distance effects emerged in any of the experiments—all of the credible intervals overlap zero. There are not even any apparent trends towards such effects, as we can see from Figure 3.4.

Table 3.3: Story-time distance analysis, probe response times: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Short	Intercept	7.31	0.04	7.23	7.40
		Medium	0.02	0.02	-0.02	0.06
		Long	0.02	0.02	-0.02	0.06
	Medium	Intercept	7.34	0.04	7.25	7.43
		Long	-0.00	0.02	-0.05	0.04
Result	Short	Intercept	7.29	0.03	7.23	7.36
		Medium	0.01	0.01	-0.02	0.04
		Long	0.00	0.01	-0.02	0.03
	Medium	Intercept	7.30	0.03	7.25	7.36
		Long	-0.01	0.01	-0.03	0.02
Explanation	Short	Intercept	7.29	0.03	7.22	7.35
		Medium	0.01	0.01	-0.02	0.04
		Long	0.02	0.02	-0.01	0.05
	Medium	Intercept	7.31	0.03	7.24	7.37
		Long	-0.00	0.02	-0.03	0.03

The data from the three experiments were then combined for analysis of coherence relation effects. The data from each experiment were pooled across story-time distances for this part of the analysis, as illustrated in Figure 3.5. Fixed effects estimates are given in Table 3.4. Although the response times for Narrations appear in Figure 3.4 to be somewhat higher than for Results and Explanations, there were no reliable effects of coherence relation in this measure.

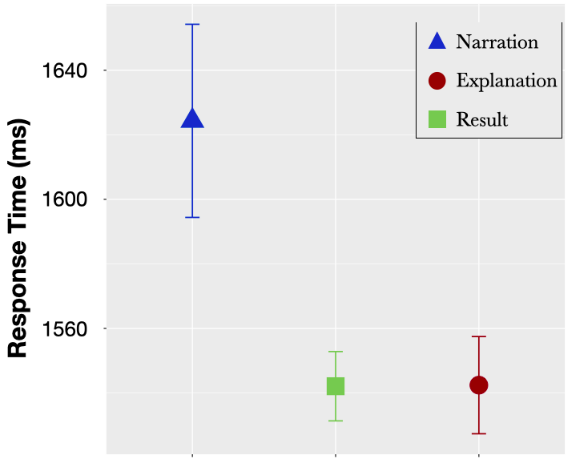


Figure 3.5: Mean probe response times, raw, in Experiments 2A-2C. Error bars show a 95% confidence interval.

Table 3.4: Coherence relation analysis, probe response times: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Intercept	7.34	0.04	7.26	7.42
	Result	-0.03	0.05	-0.13	0.08
	Explanation	-0.05	0.05	-0.14	0.04
Result	Intercept	7.31	0.03	7.26	7.37
	Explanation	-0.02	0.03	-0.09	0.05

3.4.5.3 Probe Response Accuracies

The mean proportions of correct responses to the probe questions for all three experiments are plotted in Figure 3.6.

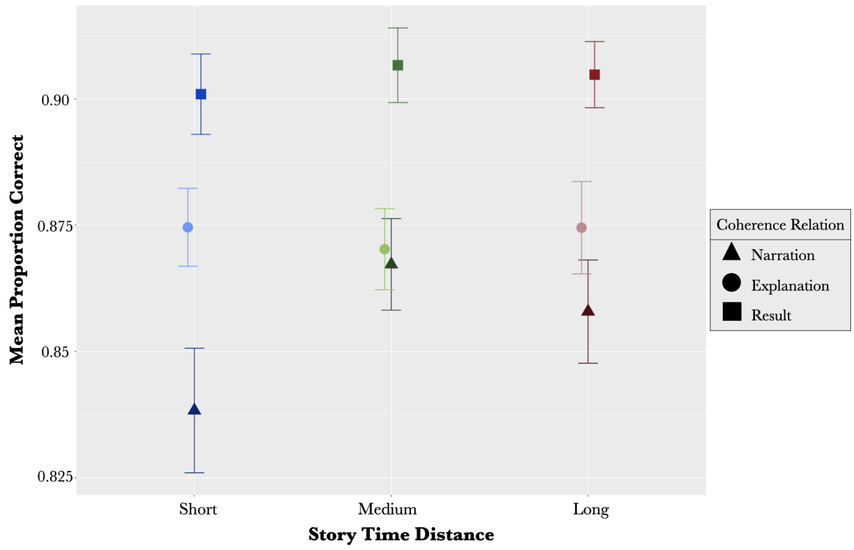


Figure 3.6: Mean probe response accuracy, in Experiments 2A-2C. Error bars show a 95% confidence interval.

The data from each experiment were first analyzed separately for story-time distance effects. Fixed effects estimates are given in Table 3.5. There were no reliable effects of story-time distance in any of the experiments.

The data from the three experiments were also combined and analyzed for effects of coherence relation. The data from each experiment were pooled across story-time distances for this part of the analysis, as illustrated in Figure 3.7. Fixed effects estimates are in Table 3.6. There were reliable effects of coherence relation: response accuracy

Table 3.5: Story-time distance analysis, probe response accuracy: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Short	Intercept	2.26	0.32	1.65	2.92
		Medium	0.32	0.30	-0.25	0.93
		Long	0.21	0.26	-0.27	0.77
	Medium	Intercept	2.43	0.32	1.84	3.11
		Long	-0.00	0.27	-0.52	0.55
Result	Short	Intercept	2.62	0.26	2.14	3.16
		Medium	0.16	0.28	-0.37	0.72
		Long	0.12	0.26	-0.38	0.68
	Medium	Intercept	2.66	0.25	2.19	3.19
		Long	0.07	0.26	-0.42	0.62
Explanation	Short	Intercept	2.69	0.34	2.04	3.38
		Medium	-0.14	0.25	-0.61	0.38
		Long	0.02	0.25	-0.46	0.54
	Medium	Intercept	2.49	0.33	1.87	3.16
		Long	0.21	0.24	-0.26	0.72

for Narration was lower than for Result and for Explanation. There was no difference between Result and Explanation, though a slight trend towards Result accuracy being higher than Explanation accuracy is visible in Figure 3.6.

Table 3.6: Coherence relation analysis, probe accuracy rate: Summary of Bayesian estimates for separate models in Experiments 2A-2C.

Reference condition		Estimate	Estimated error	2.5%	97.5%
Narration	Intercept	2.09	0.25	1.61	2.58
	Result	0.54	0.19	0.15	0.91
	Explanation	0.55	0.18	0.20	0.92
Result	Intercept	2.60	0.21	2.19	3.04
	Explanation	0.03	0.19	-0.33	0.44

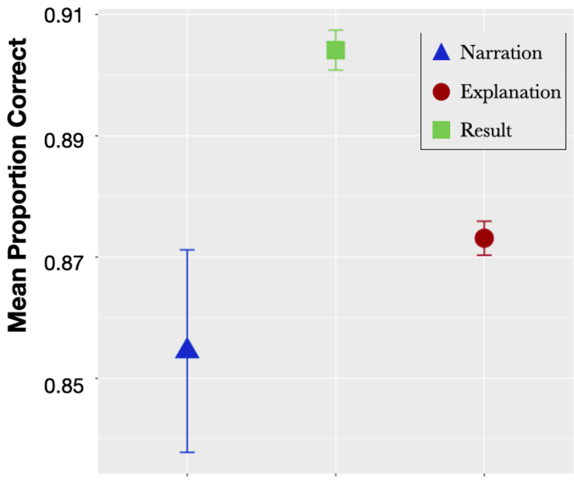


Figure 3.7: Mean probe response times, raw in Experiments 2A-2C. Error bars show a 95% confidence interval.

3.5 General Discussion

3.5.1 The temporal dimension

With respect to the Strong Iconicity vs. Scenario Model debate, these results are, unfortunately, not particularly illuminating. The predictions of these views differed with respect to story-time distance—the expected patterns are summarized in (6).

(6) a. Strong Iconicity – Reading time of $\pi 2$ and Probe response times: Short ('a moment') $<$ Medium ('ten minutes') = Long ('a day')

b. Scenario Model – Reading time of $\pi 2$ and Probe response times: Short = Medium $<$ Long

c. Strong Iconicity – Probe response accuracy: Short $>$ Medium = Long

d. Scenario Model – Probe response accuracy: Short = Medium > Long

None of the predictions of either view are borne out. For reading time of $\pi 2$, probe response times, and response accuracy, there were no reliable effects of story-time distance in any of the experiments.

The lack of story-time distance effects in the Narration experiment means that it failed to replicate either Zwaan (1996) or Rinck & Bower (2000). There are a few possible reasons for this. The first is that my experiments were not designed to be direct replications of the previous studies. Second, my experiments had more statistical power than those studies—mine had ten observations per condition per participant, compared to six for both Zwaan and Rinck & Bower, as well as more participants. Further investigation is needed, of course, but it is possible that the previous results were Type I errors. Finally, if we look more closely at the results reported by Zwaan—whose study was more directly comparable with mine—we find that the effect sizes he found were quite small. In the reading times, the short condition was faster than the medium and long conditions by only about 100 ms. For the probe response times, the short vs. medium/long difference was about 65 ms. In the raw reading times in my Narration experiment (2A), there is a pattern roughly similar to Zwaan's: short (2228 ms) vs. medium (2314 ms) vs. long (2368 ms). After correction, though, no reliable effects came out.

The lack of story-time distance effects is compatible with the predictions of SDRT as operationalized in §3.4.1.2, but as discussed in that section, it is possible to operationalize SDRT in ways that lead to different predictions. These experiments alone cannot tell us much about the validity of SDRT in this respect.

3.5.2 The causal dimension

Although the temporal dimension results are a wash, we can glean some insight into the role of causal linking in situation model construction. Strong Iconicity and the Scenario Model in their most basic forms both predict that Explanation, by virtue of entailing temporal backshift, should incur higher $\pi 2$ reading times, higher probe RTs, and lower probe recognition accuracy than the progression-entailing Narration and Result. This is not borne out in the current experiments, a result that converges with the SPR and Maze findings of Experiment 1. The lack of differences between Result (2B) and Explanation (2C), in particular, suggests that backshift is not inherently more difficult to process than progression, nor does it necessarily constitute a discontinuity in a situation model. In the raw reading times, there was a trend towards Explanations being read more slowly than the Narrations and Results, but it turned out not to be a reliable effect. Perhaps a (weak) iconicity effect does exist, but was not picked up in full-sentence reading times. Future work in the Maze or eye-tracking while reading should explore whether iconicity effects turn up in incremental processing.

Perhaps the most interesting finding is that the non-causal Narration had both a trend towards longer probe RTs and significantly lower probe recognition accuracy than the causal Result and Explanation. It suggests that Narration, compared to Result and Explanation, may constitute a discontinuity in the causal dimension. In SMT, this difference between Narration and Result is expected. That is, it is reasonable to think that encoding a causal link between $\pi 1$ and $\pi 2$ may require more reactivation of the $\pi 1$ representation than a mere progression link, so performance in memory tasks is better when there is a causal link. There is evidence in other domains that elaboration and greater representational complexity can strengthen a memory representation, thus conferring a retrieval advantage (Hofmeister (2011); Karimi (2019), a.o.).[4]

Under a view of SMT in which the temporal and causal dimensions have equal influence on situation model construction, the difference between Narrations and Explanations is unexpected. If the encoding strength/reactivation view is on the right track, though, this pattern may not be so surprising. Further, it is still possible that Explanations are more difficult than Narrations (and Results) in incremental processing because of their temporal iconicity violation. Thus, in this case, the probe performance advantage that Explanation has over Narration could be an instance of reanalysis as a form of elaboration that leads to later advantages (Arnett & Wagers, 2017). That is, if (relatively) early processing of Explanations involves reanalysis of the temporal relation, i.e., switching to backshift from first-choice progression, the reactivation this requires could boost later retrieval performance.

[4] I am extremely grateful to Nick Van Handel for suggesting this interpretation to me.

Taken together, the results of these experiments suggest that the temporal and causal dimensions are not equally influential in SMT—we observed no effects of purported temporal discontinuities, but did see what appear to be causal discontinuity effects. And, if the hint of an iconicity effect in the reading time trends can be found in future studies, it may be the case that temporal dimension effects like the story-time distance effect are modulated by the causal dimension. It is not surprising that the dimensions are not all equally influential. Previous work has shown that the spatial dimension does not seem to be monitored all the time, but the temporal and causal dimensions were previously thought to be equally influential (Zwaan et al. (1995); Therriault et al. (2006); Komeda & Kusumi (2006); Radvansky & Copeland (2010); Bailey et al. (2017), a.o.). It is possible to adjust the SMT framework a bit, to reweight the influences of the temporal and causal dimensions. However, it is not clear if this would be enough to capture the way in which the dimensions seem to interact with one another—it is a drawback of the framework that there are few explicit avenues for interdimensional contact.

Chapter 4

Making sense of the Right Frontier

4.1 Introduction

In this chapter, I turn to Segmented Discourse Representation Theory (SDRT), the linguistic coherence theory previewed in Chapter 2. In particular, I focus on the Right Frontier Constraint, a structural constraint on discourse-level attachments and pronominal anaphora resolution. I argue that the Right Frontier Constraint is the part of SDRT that can be most straightforwardly operationalized, the primary source of SDRT's empirically-testable predictions. I also propose that the reduced-coherence effects that are attributed to the RFC under SDRT are actually a conglomeration of the effects of a variety of factors that influence entity salience in discourse. In a pair of stops-making-sense experiments, I test some of SDRT's predictions stemming from the

RFC, as well as the proposal that RFC effects can be accounted for by independent factors.

I outline the SDRT framework in §4.2. In §4.3, I provide more detail of how discourse structure is constructed incrementally in the framework, and give an overview of the Right Frontier Constraint. In §4.4, I present Experiment 3A, which investigates how animacy and the RFC interact. The findings here suggest that animacy hierarchy effects are partly responsible for what we would simply classify as RFC effects under SDRT. In §4.5, I present Experiment 3B, which investigates how event structure and the RFC interact. In this experiment, we found RFC-type effects, but the specific manipulation we used did not support the view that event structure effects are partly responsible for RFC effects—at least not as straightforwardly as the animacy results. Additionally, the results of these studies suggest that the stops-making-sense paradigm, which has previously been used only to study sentence processing phenomena, is also a valid methodology for discourse-level investigations.

4.2 Basics of SDRT

4.2.1 The roots of coherence theory

A fundamental question in the study of discourse is this: What makes a sequence of clauses or sentences seem like a coherent discourse? Examining even tiny discourses, like those in (1), shows that purely linguistic information cannot completely answer that question. To start, (1-a) does not form a coherent discourse–it is merely a pair of adjacent, unrelated sentences. In (1-b), the shift from 'Art' to 'she' in the second sentence makes for a more coherent sequence of sentences, provided that one infers that 'she' refers to 'Melanie.' Still, (1-b) is not a very natural discourse. Finally, (1-c) forms a reasonably coherent, natural discourse. One can pretty straightforwardly infer that Melanie went to Istanbul *because* she has family there.

(1) a. Melanie took a train from Paris to Istanbul. Art likes spinach.
  b. Melanie took a train from Paris to Istanbul. She likes spinach.
  c. Melanie took a train from Paris to Istanbul. She has family there.

Crucially, the inference for (1-c) is possible even in the absence of an explicit linguistic cue of the causal relationship. And, in comparison, (1-b) is not as coherent as (1-c), despite being better than (1-a), presumably because of the cross-sentential anaphoric link between 'Melanie' and 'she.' It seems, then, that certain linguistic information, such as pronominal anaphora, can increase coherence, but is not its main driver. Rather, coherence is determined by inference processes which stem from our world knowledge, and are influenced by linguistic information. If one can fairly readily infer that there is a plausible relationship between two described eventualities, then those descriptions can be interpreted as components of a single coherent discourse.

Beginning with Hobbs (1979), formal linguistic and NLP research have worked to formally characterize how the relationship between world knowledge-based inference and language determines the well-formedness of discourses. The various theories in this domain can be broadly termed 'coherence theory.' Different coherence theories disagree on how highly structured discourse is. However, they are largely in agreement on certain other foundational assumptions and basic building blocks. The building blocks—coherence relations—are inferred links between the events and states described in a discourse. The idea is that, in combination with our world knowledge and an essentially Gricean inferential system, we need only a small set of coherence relations to produce and interpret coherent discourses of any size.

In Hobbs' (1979, 1985) foundational coherence theory works, comprehenders establish local coherence between adjacent clauses by inferring coherence relations that link them. On his view, a discourse has global coherence if each part of it fits into an identifiable 'plan,' e.g., a task model or a speaker's conversational plan. However, it is possible for comprehenders to judge whether a discourse is coherent even if they are not familiar with the 'plan' it is meant to follow. More recent coherence theories have decentralized the notion of 'plans' in creating and assessing coherence, and define global coherence differently, as we will shortly see. Two major contemporary coherence theories, developed by Asher and Lascarides (2003) and Kehler (2002), hew fairly closely to Hobbs' conception of local coherence, i.e., his small set of coherence relations.

Kehler (2002) sticks more closely to Hobbs' definitions and categorizations of coherence relations than do Asher and Lascarides (2003). However, by his own admission, Kehler does not make claims about the mechanics of the inference process, nor about how comprehenders ultimately choose a discourse relation. Additionally, there is no notion of hierarchical discourse structure under his view. Because Kehler makes so few claims about how discourses are either built or comprehended, it is difficult to see whether his version of coherence theory is falsifiable.

The pragmatic inferencing machinery of SDRT draws heavily from Hobbs. In contrast to other Hobbs-descended theories, e.g., Kehler (2002), SDRT also has significant, explicit semantic machinery. This is drawn from Discourse Representation Theory (DRT), a dynamic semantics pioneered by Kamp (1981) (see also Kamp & Reyle 1993). The core idea of the original DRT is that comprehenders incrementally build a mental representation, called a discourse representation structure (DRS), as they progress through a discourse. DRSs, which are graphically represented as in (2), include a set of discourse referents, and a set of conditions that encode information about those discourse referents.

(2) Peter owns a Peugeot. It doesn't run.

$$
\boxed{\begin{array}{c}
\mathrm{x\ y\ z} \\
\mathrm{Peter(x)} \\
\mathrm{Peugeot(y)} \\
\mathrm{x\ owns\ y} \\
\mathrm{z = y} \\
\neg\!: \boxed{\mathrm{z\ runs}}
\end{array}}
$$

DRT relies heavily on linguistic cues to establish temporal relations in discourse. As such, it incorrectly predicts that (3-a) and (3-b) must have the same temporal relation—progression—because they have the same tense and aktionsart. While progression is indeed the most natural interpretation of (3-a), this is not true of (3-b). The most readily available interpretation of (3-b) is that David's falling was caused by Alexis' pushing, so the temporal relation between the sentences must be a backshift. In DRT, however, this discourse may only be interpreted as a progression.

(3) a. David fell. Alexis helped him up.

b. David fell. Alexis pushed him.

So, while DRT is semantically sophisticated, its reliance on linguistic cues and relatively impoverished pragmatic machinery limit its empirical coverage. Hobbs' and Kehler's coherence theories, on the other hand, are relatively better-articulated with respect to pragmatics/inferential reasoning, but semantically underspecified. Lascarides and Asher's approach is to combine DRT, which provides firm semantic grounding, and Hobbsian coherence theory, which provides explicit pragmatic constraints on discourse structure, into SDRT.

4.2.2 Essential SDRT relations

As in classic DRT, individual clauses and sentences in SDRT are represented as DRSs. Discourses are represented a Segmented DRSs, which consist of a set of DRSs and a set of conditions (e.g., coherence relations) on those DRSs. SDRSs are built up through the incremental construction of DRSs for individual clauses, then the incremental attachment of those DRSs to one another. The core discourse relations for narrative discourses—as opposed to dialogues—are defined and exemplified in (4)-(8).

(4) a. Narration: The event described in $\pi 2$ is a consequence of (but not strictly speaking caused by) the event described in $\pi 1$.

b. Jendaya watered and weeded her garden. She pruned her Japanese maple.

(5) a. Result: The event described in $\pi 1$ caused the event or state described in $\pi 2$.

b. Jendaya watered and weeded her garden. Her beloved plants began to flourish.

(6) a. Explanation: The event described in $\pi 2$ explains why $\pi 1$'s event happened (perhaps by causing it).

b. Jendaya watered and weeded her garden. She noticed it was looking dry and overgrown.

(7) a. Elaboration: $\pi 2$'s event is part of $\pi 1$'s (perhaps by being in the preparatory phase, in the sense of Moens & Steedman 1988).

b. Jendaya watered and weeded her garden. She was careful not to overwater the succulents.

(8) a. Background: The state described in $\pi 2$ is the 'backdrop' or circumstances under which the event in $\pi 1$ occurred (no causal connection but the event and state temporally overlap).

b. Jendaya watered and weeded her garden. It took up half of the backyard.

These relations can be grouped into two categories—coordinating and subordinating—based on the types of discourse structures they produce. Narration, Result, and Background are all coordinating relations, in that they extend the main spine of a discourse structure. On the other hand, Elaboration and Explanation are subordinating—they do not advance a narrative discourse, but provide further information about an eventuality that has already been introduced. As we will see in §4.3, the coordination/subordination distinction plays an important role on the Right Frontier.

Asher & Lascarides's goal is a computationally tractable theory of discourse structure, so they need some way to constrain the ways in which non-linguistic factors such as world knowledge or 'agents' cognitive states' may influence decisions about discourse structure. To that end, they propose a nonmonotonic glue logic that admits enough world knowledge to narrow down the set of possible coherence relations at a given point, but does not require full interpretation in a strictly incremental way.

The nonmonotonicity of the glue logic allows Asher & Lascarides to capture the ambiguity of sequences like that in (9)—a backshift relation may be inferred (and indeed preferred) based on world knowledge, though a progression relation is also available.

(9) John went to jail. He embezzled the company funds (181).

4.3 Constructing local and global coherence

4.3.1 Basics of SDRS construction

4.3.1.1 The smallest discourses

First, we will build an SDRS for the discourse in (10).

(10) a. [$\pi 1$] Jendaya watered and weeded her vegetable garden.

b. [$\pi 2$] Then, she pruned her Japanese maple.

We can infer that this is an instance of Narration—the sentences are in a progression relation, and are not causally linked. The sentence-by-sentence SDRS construction is schematized in (11).

(11) a. $[\pi 1]$

b. $[\pi 1]$ —— $[\pi 2]$

Narration extends the discourse, so it is a Coordinating relation. As such, it is represented horizontally in the SDRS. This distinguishes it from Subordinating relations such as Explanation, which we can see an instance of in (12).

(12) a. [π1] Tristan watered and weeded his vegetable garden.

b. [π2] He noticed that it was looking dry and overgrown.

Explanation does not advance the discourse. Rather, it fills in new detail about the first sentence. As such, it is a Subordinating relation, and is represented vertically in an SDRS, as in (13).

(13) a. $[\pi 1]$

$$\begin{array}{c} \pi 1 \\ | \\ \pi 2 \end{array}$$

b.

4.3.1.2 Building larger discourses

An example of the incremental SDRS construction for the discourse in (14) is shown in (15). The coherence relation between (14-a) and (14-b) is Narration, as is the relation between (14-b) and (14-c). As each clause is introduced, the comprehender must 'attach' the new clause to the preceding clause by inferring a coherence relation between them.

(14) a. $\pi 1$: Jendaya took all sorts of cooking classes during her high school years.

b. $\pi 2$: Then she became a line cook at a popular restaurant in her hometown.

c. $\pi 3$: And, just last year, she graduated from Le Cordon Bleu in Paris.

(15) a. $\pi 1$

b. $\pi 1 — \pi 2$

c. $\pi 1 — \pi 2 — \pi 3$

This discourse is fully locally coherent. That is, each clause is attached in the discourse structure to the clause that immediately precedes it in the linear order of the discourse. However, discourse relations also seem to arise between non-adjacent sentences. For example, the discourse in (16) (adapted from Polanyi (1988)) is the same as that in (14), but with an additional sentence at the beginning. The relationship between (16-a) and (16-b) is not one of progression–in fact, (16-b) represents a step backwards in time from (16-a). Instead, (16-b) provides some justification for the proposition in (16-a). The local progression/Narration relations between (16-b) and (16-c), and then between (16-c) and (16-d), are still present, but these are not the only relations that (16-c) and (16-d) are involved in. Like (16-b), they both support the claim in (16-a)–they are each interpreted as being in the same evidentiary discourse relation with (16-a) as (16-b) is, though they are not immediately adjacent to (16-a).

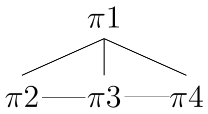
(16) a. $\pi 1$: Jendaya is an experienced chef.

b. $\pi 2$: She took all sorts of cooking classes during her high school years.

c. $\pi 3$: Then she became a line cook at a popular restaurant in her hometown.

d. $\pi 4$: And, just last year, she graduated from Le Cordon Bleu in Paris.

To account for nonlocal relations such as these, Polanyi (1988) and Grosz and Sidner (1986)—whom Asher & Lascarides follow—posit hierarchical discourse structure. In these theories, a clause or sentence $\pi 2$ is held to be embedded by another clause/sentence $\pi 1$ at the discourse-structural level if, for example, the propositional content of $\pi 2$ somehow expands on the propositional content of $\pi 1$. For instance, in the discourse in (16), the sentences in (16-b)-(16-d) are embedded under the one in (16-a), because they all expand on the propositional content of $\pi 1$ by providing justification for it. In SDRT, these are Elaborations. Besides the Elaboration links, the clauses in (16-b)-(16-d) are linked by Narration, just as they were in (14). The complete structure of the discourses in (16) can be represented as in (17). This discourse, despite having nonlocal attachments, is also fully locally coherent, since the clauses that are attached non-locally ($\pi 3$ and $\pi 4$) are also attached locally.

(17) 

Complete local coherence is not a necessary condition of a globally-coherent discourse (i.e., one that we judge to form a single integrated unit). Consider the example in (18):

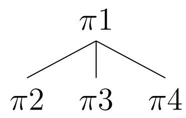
(18) a. $\pi 1$: Aislinn adopted a dog from her local animal shelter.

b. $\pi 2$: He's a pit bull mix with a bluish-gray coat.

c. $\pi 3$: He's pretty old already, and mostly likes to lie around and get belly rubs.

d. $\pi 4$: His name is Mortimer.

This is our first example of a globally coherent discourse that is not fully locally coherent. The clauses in (18-b), (18-c), and (18-d) are each in an Elaboration with (i.e., are subordinated to) (18-a)–they each provide some detail about Aislinn's dog. However, unlike the string of Elaborations in (16), these Elaborations are not also in local Narration relations with one another, as they simply describe concurrent, independent states that Aislinn's dog is in. As such, they are not in any other local coherence relation either. This discourse can be represented as in (19).

(19) 

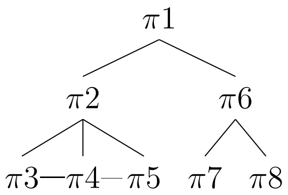
Multiple embedding—another form of local incoherence—is also possible, as illustrated in the discourse in (20) (also adapted from Polanyi 1988). The sentence in (20-b) dominates those in (20-c)-(20-e), as it does in (16), and in this case it is itself dominated by (20-a). Proceeding from (20-e), we can see that the proposition in (20-f) cannot be readily interpreted as evidence that Jendaya is a great cook, since it is not about Jendaya at all. It can, however, be easily interpreted as support for (20-a). In terms of the overall discourse structure, this means that the embedded structure dominated by (20-b) ends with (20-e). The sentence in (20-f) is dominated only by (20-a). Finally, the sentences in (20-f) and (20-f) are interpreted as Elaborations of (20-f); they are embedded by (20-f). The structure of (20) is shown in (21).

(20)
a. $\pi 1$: Lucille's housemates are quite accomplished.
b. $\pi 2$: Jendaya is an experienced chef.
c. $\pi 3$: She took all sorts of cooking classes during her high school years.
d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.
e. $\pi 5$: Just last year, she graduated from Le Cordon Bleu in Paris.

f. $\pi 6$: Marceline is a brilliant musician.

g. $\pi 7$: She plays about six different instruments, all of them really well.

h. $\pi 8$: She works with a whole bunch of famous recording artists.

(21) 

Like the discourse in (18), the discourse in (20) is interpretable as a globally coherent discourse, although it is not fully locally coherent. For instance, the sentences $\pi 5$ (20-e) and $\pi 6$ (20-f), though adjacent in the discourse, are not directly linked by any discourse relation. In fact, they're not even at the same level of embedding: $\pi 5$ is subordinated to $\pi 2$, which is itself subordinated to $\pi 1$, while $\pi 6$ is subordinated to $\pi 1$.

4.3.2 The Right Frontier Constraint

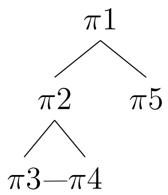
However, a discourse can't have breaks in local coherence just anywhere, and still be globally coherent. Compare (20) and (22): they contain exactly the same propositional content, and differ only in the relative linear order of the two adjacent sentences expressing, respectively, that Jendaya graduated from Le Cordon Bleu and that Marceline is a brilliant musician. In both versions of the discourse, there is no discourse relation directly linking these two propositions. However, as (22) shows, changing their relative linear ordering results in an incoherent discourse. If the antecedent of *she* in (22-f)/$\pi 6$ is taken to be Marceline, the discourse is rendered incoherent—or at least less coherent than (20)—because graduating from Le Cordon Bleu doesn't seem to have anything to do with being a brilliant musician—no discourse relation can be readily inferred to link (22-e) and (22-f). If the antecedent of *she* is instead Jendaya, the discourse is still incoherent, this time because apparently one can't simply resume talking about Jendaya right after starting to talk about Marceline.[1]

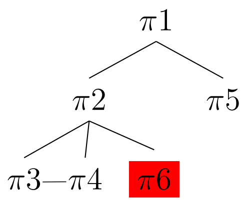
(22)
a. $\pi 1$: Lucille's housemates are quite accomplished.
b. $\pi 2$: Jendaya is an experienced chef.
c. $\pi 3$: She took all sorts of cooking classes during her high school years.
d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.
e. $\pi 6$: Marceline is a brilliant musician.
f. $\pi 5$: ??Just last year, she$_{m/j}$ graduated from Le Cordon Bleu in Paris.

The relative incoherence of (22)—in particular when the *she* in (22-f) corefers with *Jendaya*—shows us that a theory of discourse coherence that involves hierarchical structure must include restrictions on how and where an incoming clause can be attached to the existing discourse structure. Let's look at a partial structure of (22) to better understand the problem. The schema in (22-a) illustrates the state of the discourse structure just after pi-5 has been attached. The sentence in $\pi 5$ is attached as a

[1] Further, replacing the *she* in (22-f) with *Jendaya* in an attempt to avoid the pitfalls of ambiguous pronominal reference does not improve matters.

subordinate to $\pi 1$, since it is most straightforwardly inferred to be an Elaboration of $\pi 1$. It doesn't make nearly as much sense to infer that it is an Elaboration of $\pi 4$, nor that it is temporally dependent on $\pi 4$ (as a Narration would be). Under an interpretation of (22-f)/$\pi 6$ in which she refers to Jendaya, it would be necessary to attach $\pi 6$ as subordinate of $\pi 2$, since it is another Elaboration of $\pi 2$. The intended discourse structure for this interpretation is in (22-b), with the relevant (attempted) attachment in red.

(23) a. 

b. 

Hierarchical theories of discourse structure account for the distinction between (20) and (22) by rendering the attachment of $\pi 6$ in (23-b) illicit via a constraint on the dynamics of attachment/non-local attachment. This constraint—the Right Frontier Constraint (RFC)—was originated by Polanyi (1988), and adopted by Lascarides and Asher. The constraint dictates which node(s) in an existing discourse structure are accessible for attachment. An informal definition of the RFC is given in (31).[2]

[2]The full definition of the RFC for attachment points is as follows:

Suppose that a constituent β is to be attached to a constituent in the SDRS $\langle A, F, LAST\rangle$ with a discourse relation other than Parallel or Contrast. Then the available attachment points for β are:

(24) For a clause that is to be attached to an existing discourse structure, the accessible attachment points are:

a. The most recently attached node,

b. Any node that immediately dominates the most recent node,

c. Any node that transitively dominates the most recent node.

In sum, the nodes described in this definition constitute the Right Frontier of a discourse structure, and are the only nodes accessible for attachment. The Right Frontier changes with each new clause that is integrated into the structure.

Given this constraint, we can accurately capture the coherence of (20) and the incoherence of (22). Let's first examine discourse (20)—repeated in (32)—partway through its construction. Assume that the discourse has proceeded as far as (32-e). The discourse structure at this point is as in (33-b).

(25) a. $\pi 1$: Lucille's housemates are quite accomplished.

b. $\pi 2$: Jendaya is an experienced chef.

c. $\pi 3$: She took all sorts of cooking classes during her high school years.

d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.

1. The label α = LAST;

2. Any label γ such that $\alpha < \gamma$, or:

(a) *i-outscopes*(γ, α) (i.e., $R(\delta, \alpha)$ or $R(\alpha, \delta)$ is a conjunct in $F(\gamma)$ for some *R* and some δ), or

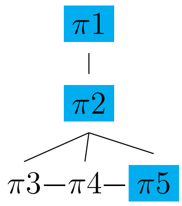
(b) $R(\gamma, \alpha)$ is a conjunt in $F(\lambda)$ for some label λ, where *R* is a subordinating discourse relation (like Elaboration or Explanation)

3. Transitive Closure: Any label γ that dominates α through a sequence of labels $\gamma_1, ..., \gamma_n$ such that $\alpha < \gamma_1, \gamma_1 < \gamma_2, ..., \gamma_n < \gamma$.

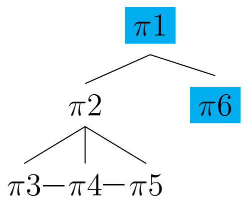
e. $\pi 5$: Just last year, she graduated from Le Cordon Bleu in Paris.

f. $\pi 6$: Marceline is a brilliant musician.

(26) a.



b.



In the structure in (33-b), three nodes are on the Right Frontier:

1. $\pi 5$ because it is the most recently attached node,

2. $\pi 2$ because it immediately dominates $\pi 5$, and

3. $\pi 1$ because it transitively dominates $\pi 5$.

Thus, it is possible under the RFC to attach $\pi 6$ (sentence (32-f)) directly to $\pi 1$. This attachment is shown in (33-b).

Now let's examine the structure of (22) at the crucial juncture. Assume the discourse has gone as far as (34-d). The discourse structure to that point is in (35-a), in which the Right Frontier nodes are $\pi 1$, $\pi 2$, and $\pi 4$. When the discourse progresses to $\pi 5$, that sentence gets attached to $\pi 1$, which is on the Right Frontier and therefore a licit attachment site. This yields the structure in (35-b).

(27) a. $\pi 1$: Lucille's housemates are quite accomplished.

b. $\pi 2$: Jendaya is an experienced chef.

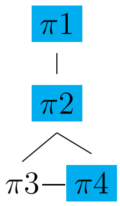
c. $\pi 3$: She took all sorts of cooking classes during her high school years.

d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.

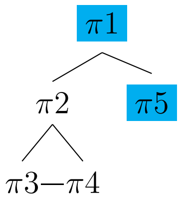
e. $\pi 6$: Marceline is a brilliant musician.

f. $\pi 5$: ??Just last year, she$_{m/j}$ graduated from Le Cordon Bleu in Paris.

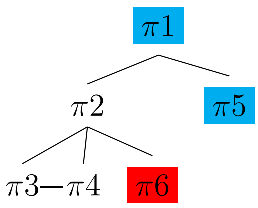
(28) a.



b.



c.



Once $\pi 5$ is attached, the Right Frontier is updated. The new Right Frontier contains only two nodes: $\pi 5$, since it's the most recently attached, and $\pi 1$, since it immediately dominates $\pi 5$. Crucially, the subordinated structure that begins with $\pi 2$ is no longer on the Right Frontier–it is sealed off from further attachments. As a result, when the discourse progresses to $\pi 6$, that sentence cannot be attached as an Elaboration of $\pi 2$, even though that is the only relation it can be readily inferred to be in. Although it is structurally possible to attach it to $\pi 1$ or $\pi 5$, no discourse relation can plausibly be inferred between $\pi 6$ and either of those.[3] Judgments of the (in)coherence of discourse (22) and others of the same type will surely vary, not just between readers, but quite possibly also for a single reader across time, depending on how willing or able they are to draw the necessary coherence inferences.

Asher & Lascarides (2003) and others argue that the combination of the RFC and the appropriate semantics of discourse relations accounts not only for constraints on discourse-structural attachment, but also for constraints on pronominal anaphora. Roughly, a pronoun in the incoming sentence can only find its antecedent in a sentence that is on the Right Frontier. In other words, if two sentences can be attached in the discourse structure, pronominal links between them are licit; otherwise, such links are illicit.

Let's examine another discourse, this time from Asher and Lascarides, in which the purported influence of the RFC on pronominal anaphora is fairly clear. The discourse is in (36). Its structure after attaching (36-e) and before attaching (or attempting to attach) (36-f) is shown in (37).

(29) a. $\pi 1$: Mahina had a great evening last night.

b. $\pi 2$: She had a great meal.

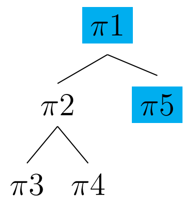
[3]And if the *she* in $\pi 6$ is replaced with *Jendaya*, it becomes easier/possible to infer that $\pi 6$ is an Elaboration of $\pi 1$, and so to attach it to $\pi 1$ via Subordination. However, it feels odd to create a second subordinate structure about Jendaya's cooking bona fides right after creating a subordinate structure about Marceline.

c. $\pi 3$: She ate salmon.

d. $\pi 4$: She devoured lots of cheese.

e. $\pi 6$: She then won a dancing competition.

f. $\pi 5$: #It/#The salmon was a beautiful pink.

(30) 

At this stage, $\pi 5$ is on the Right Frontier because it is the newest node, and $\pi 1$ is on the right frontier because it immediately dominates $\pi 5$. Though it's structurally possible to attach $\pi 6$ to $\pi 5$, *It/The salmon was a beautiful pink* can't be reasonably inferred to be in any direct discourse relation with *She then won a dancing competition*. As for the other open attachment site, $\pi 1$, it may be possible to infer a pretty flimsy Subordination link with the version of $\pi 6$ that begins with *the salmon*. However, the version beginning with *it* really cannot be linked to $\pi 1$, since in that case the only possible antecedents of *it* are *great evening* and *last night*, neither of which can plausibly be described as being *a beautiful pink*.

The optimal intended attachment for $\pi 6$ is probably a Subordination from $\pi 3$, as an Elaboration about the salmon, though it could also plausibly be a Subordination from $\pi 2$. However, once $\pi 5$ is attached as it is in (36), neither $\pi 2$ nor $\pi 3$ is on the Right Frontier, so none of the entities they describe are accessible as referents for pronominal anaphora.

In this section, I've gone to some lengths in discussing the Right Frontier Constraint in order to show that it can generate clear, straightforwardly testable predictions about both offline acceptability and online accessibility in discourse. In fact, most of the predictions that SDRT can generate stem from the RFC. The RFC is essentially a 'grammatical' constraint on the construction of discourse structure, however, and it is an open question whether it really ought to be encoded in this way, or whether effects that are attributed to it can actually be boiled down to independently-attested effects, such as animacy (hierarchy) effects. In the rest of this chapter, I present experiments that begin to address this question.

As I argued in §4.3, the bulk of the testable predictions of SDRT come from the Right Frontier Constraint, because it is argued to constrain both pronominal anaphora resolution and discourse-structural attachment. At its core, the RFC is a grammatical constraint on discourse-structure building. WHierarchical theories of discourse structure account for the distinction between (20) and (22) by rendering the attachment of $\pi 6$ in (23-b) illicit via a constraint on the dynamics of attachment/non-local attachment. This constraint—the Right Frontier Constraint (RFC)—was originated by Polanyi (1988), and adopted by Lascarides and Asher. The constraint dictates which node(s) in an existing discourse structure are accessible for attachment. An informal definition of the RFC is given in (31).

(31) For a clause that is to be attached to an existing discourse structure, the accessible attachment points are:

a. The most recently attached node,

b. Any node that immediately dominates the most recent node,

c. Any node that transitively dominates the most recent node.

In sum, the nodes described in this definition constitute the Right Frontier of a discourse structure, and are the only nodes accessible for attachment. The Right Frontier changes with each new clause that is integrated into the structure.

Given this constraint, we can accurately capture the coherence of (20) and the incoherence of (22). Let's first examine discourse (20)—repeated in (32)—partway through its construction. Assume that the discourse has proceeded as far as (32-e). The discourse structure at this point is as in (33-b).

(32) a. $\pi 1$: Lucille's housemates are quite accomplished.

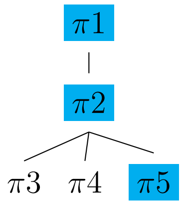
b. $\pi 2$: Jendaya is an experienced chef.

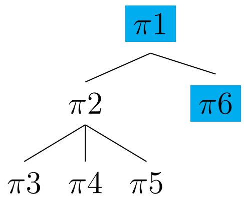
c. $\pi 3$: She took all sorts of cooking classes during her high school years.

d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.

e. $\pi 5$: Just last year, she graduated from Le Cordon Bleu in Paris.

f. $\pi 6$: Marceline is a brilliant musician.

(33) a. 

b. 

In the structure in (33-b), three nodes are on the Right Frontier:

1. $\pi 5$ because it is the most recently attached node,
2. $\pi 2$ because it immediately dominates $\pi 5$, and
3. $\pi 1$ because it transitively dominates $\pi 5$.

Thus, it is possible under the RFC to attach $\pi 6$ (sentence (32-f)) directly to $\pi 1$. This attachment is shown in (33-b).

Now let's examine the structure of (22) at the crucial juncture. Assume the discourse has gone as far as (34-d). The discourse structure to that point is in (35-a), in which the Right Frontier nodes are $\pi 1$, $\pi 2$, and $\pi 4$. When the discourse progresses to $\pi 5$, that sentence gets attached to $\pi 1$, which is on the Right Frontier and therefore a licit attachment site. This yields the structure in (35-b).

(34) a. $\pi 1$: Lucille's housemates are quite accomplished.

b. $\pi 2$: Jendaya is an experienced chef.

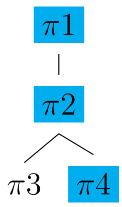
c. $\pi 3$: She took all sorts of cooking classes during her high school years.

d. $\pi 4$: Then she was a line cook at a popular restaurant in her hometown.

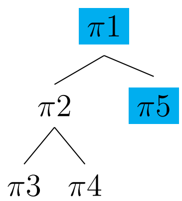
e. $\pi 6$: Marceline is a brilliant musician.

f. $\pi 5$: ??Just last year, she$_{m/j}$ graduated from Le Cordon Bleu in Paris.

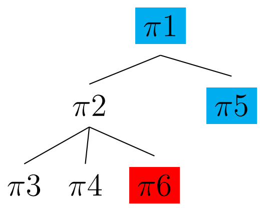
(35) a.



b.



c.



Once $\pi 5$ is attached, the Right Frontier is updated. The new Right Frontier contains only two nodes: $\pi 5$, since it's the most recently attached, and $\pi 1$, since it immediately dominates $\pi 5$. Crucially, the subordinated structure that begins with $\pi 2$ is no longer on the Right Frontier–it is sealed off from further attachments. As a result, when the discourse progresses to $\pi 6$, that sentence cannot be attached as an Elaboration of $\pi 2$, even though that is the only relation it can be readily inferred to be in. Although it is structurally possible to attach it to $\pi 1$ or $\pi 5$, no discourse relation can plausibly be inferred between $\pi 6$ and either of those.[4] Judgments of the (in)coherence of discourse (22) and others of the same type will surely vary, not just between readers, but quite possibly also for a single reader across time, depending on how willing or able they are to draw the necessary coherence inferences.

Asher & Lascarides (2003) and others argue that the combination of the RFC and the appropriate semantics of discourse relations accounts not only for constraints on discourse-structural attachment, but also for constraints on pronominal anaphora. Roughly, a pronoun in the incoming sentence can only find its antecedent in a sentence that is on the Right Frontier. In other words, if two sentences can be attached in the discourse structure, pronominal links between them are licit; otherwise, such links are illicit.

Let's examine another discourse, this time from Asher and Lascarides, in which the purported influence of the RFC on pronominal anaphora is fairly clear. The discourse is in (36). Its structure after attaching (36-e) and before attaching (or attempting to attach) (36-f) is shown in (37).

(36) a. $\pi 1$: Mahina had a great evening last night.

b. $\pi 2$: She had a great meal.

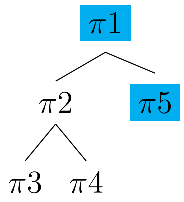
c. $\pi 3$: She ate salmon.

[4] And if the *she* in $\pi 6$ is replaced with *Jendaya*, it becomes easier/possible to infer that $\pi 6$ is an Elaboration of $\pi 1$, and so to attach it to $\pi 1$ via Subordination. However, it feels odd to create a second subordinate structure about Jendaya's cooking bona fides right after creating a subordinate structure about Marceline.

d. $\pi 4$: She devoured lots of cheese.

e. $\pi 6$: She then won a dancing competition.

f. $\pi 5$: #It/#The salmon was a beautiful pink.

(37) 

At this stage, $\pi 5$ is on the Right Frontier because it is the newest node, and $\pi 1$ is on the right frontier because it immediately dominates $\pi 5$. Though it's structurally possible to attach $\pi 6$ to $\pi 5$, *It/The salmon was a beautiful pink* can't be reasonably inferred to be in any direct discourse relation with *She then won a dancing competition*. As for the other open attachment site, $\pi 1$, it may be possible to infer a pretty flimsy Subordination link with the version of $\pi 6$ that begins with *the salmon*. However, the version beginning with *it* really cannot be linked to $\pi 1$, since in that case the only possible antecedents of *it* are *great evening* and *last night*, neither of which can plausibly be described as being *a beautiful pink*.

The optimal intended attachment for $\pi 6$ is probably a Subordination from $\pi 3$, as an Elaboration about the salmon, though it could also plausibly be a Subordination from $\pi 2$. However, once $\pi 5$ is attached as it is in (36), neither $\pi 2$ nor $\pi 3$ is on the Right Frontier, so none of the entities they describe are accessible as referents for pronominal anaphora.

In this section, I've gone to some lengths in discussing the Right Frontier Constraint in order to show that it can generate clear, straightforwardly testable predictions about both offline acceptability and online accessibility in discourse. In fact, most of the predictions that SDRT can generate stem from the RFC, because it constrains both pronominal anaphora resolution and discourse-structural attachment. There are certainly empirically-observable limits on pronominal anaphora resolution and clause-level discourse attachment. However, at its core, the RFC is a grammatical constraint on discourse-structure building, and it is an open question whether the limits on anaphora and attachment in discourse should be encoded in the 'grammar' of discourse-structure building, or are better conceived of as a constellation of effects of various linguistic and world-knowledge factors. In the rest of this chapter, I present two studies that investigate how two factors—animacy and event structure—modulate pronominal anaphora resolution. In Experiment 3A (§4.4), I find evidence of RFC-type effects, as well as evidence to suggest that animacy is partly responsible for those effects. In Experiment 3B (§4.5), RFC-type effects also emerge, while the role of event structure in them seems to be somewhat more complicated than the role of animacy.

4.4 Experiment 3A: Animacy and the RFC

4.4.1 Design and Predictions

This experiment investigated the effects of animacy on long-distance pronominal anaphora resolution. Its goal was to probe whether reduced-acceptability effects that SDRT attributes to the RFC can instead be at least partially attributed to animacy hierarchy effects. Animacy is often held to be organized in an implicational hierarchy, as in (38) (Smith-Stark, 1974; Silverstein, 1976).

(38) human > animate > inanimate

This hierarchy has been hypothesized to be correspond to the 'salience' or 'accessibility' of an entity (Zaenen et al., 2004; a.o.)—the higher on the hierarchy, the more salient/accessible an entity ought to be. This in turn corresponds to how readily an entity can be identified as a pronominal antecedent. My hypothesis is that the RFC as a constraint on pronominal anaphora indirectly addresses the effects of salience on pronominal anaphora. Thus, individual factors that seem to modulate salience, a notoriously ill-defined concept, may have isolable effects on pronominal anaphora resolution in discourse. In the case of animacy, a human entity is expected to be a more accessible pronominal referent than an inanimate entity.

This experiment has a 2x2 design, crossing Animacy {Animate, Inanimate} with

RFC {RFC-observing (i.e., No violation), RFC-violating}. The RFC-violating conditions had a pronoun whose intended referent was not on the Right Frontier, while the RFC-observing conditions had a definite NP instead of a pronoun.[5]

The experiment employs the stops-making-sense (SMS) task (Boland, Tanenhaus, & Garnsey, 1990; Boland, Tanenhaus, Garnsey, & Carlson, 1995), a methodology originally used to study syntactic processing that combines self-paced reading with acceptability judgment. In the classic SMS paradigm, comprehenders advance through a sentence word-by-word or phrase-by-phrase. For each word or phrase, they must decide whether or not the sentence still makes sense, and indicate their judgment by button press.

The current experiment is what I believe to be a novel attempt at scaling the SMS task up to the discourse level. The linking hypothesis is that sentences containing pronouns that cannot be coherently resolved (i.e., that do not have a plausible and accessible referent in the discourse) will be more likely to be rejected than both sentences containing resolvable pronouns and sentences containing no pronouns.

If the RFC is a blanket 'grammatical' restriction on pronominal anaphora in discourse, all RFC-violating pronoun resolutions should be equally likely to be rejected as nonsensical, regardless of the animacy of the intended referent. They should also all

[5]Labeling the definite NP conditions 'RFC-observing' is a bit of an oversimplification. The coreferential interpretation of these conditions actually does violate the RFC in its strictest sense. To get around this puzzle, Asher (2008) proposes that the RFC is not a constraint on definites, and that definites find their antecedents/coreferents in discourse via inferred topics that get added into SDRSs.

be rejected at fairly high rates. If, instead, RFC effects are an amalgam of the effects of animacy and other factors contributing to salience, then RFC-violating pronoun resolutions with a human antecedent should be less likely to be rejected than those with an inanimate antecedent. Because I hypothesize that animacy effects constitute only a part of salience effects, I also predict that a discourse with an RFC violation involving a human antecedent will still be rejected at a higher rate than any discourse with no RFC violations.

4.4.2 Materials

4.4.2.1 Discourses

This experiment included 40 experimental stimuli and 40 fillers. All stimuli were four-sentence narrative discourses. A sample item is shown in (39). The relevant pronouns and their intended antecedents are in bold. The complete materials are provided in Appendix D.1. Per the predictions outlined in the previous section, I predict that (39-a) and (39-c) will not be rejected at all, since they have no RFC violations and are otherwise designed to be coherent throughout. The condition illustrated in (39-b)—an RFC violation with a human antecedent, is predicted to be rejected at the sentence containing the pronoun (i.e., the third sentence) at a nonzero rate. I predict that the (39-d) condition—an RFC violation with inanimate antecedent—will be rejected at the third sentence at a higher rate than (39-b).

(39) a. **Animate, RFC-observing**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The man** fell over. Jessie circled back to investigate what had happened.

b. **Animate, RFC-violating**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **He** fell over. Jessie circled back to investigate what had happened.

c. **Inanimate, RFC-observing**: Jessie saw **a rickety mailbox** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The mailbox** fell over. Jessie circled back to investigate what had happened.

d. **Inanimate, RFC-violating**: Jessie saw **a rickety mailbox** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **It** fell over. Jessie circled back to investigate what had happened.

Comprehension question (easy): Was Jessie on a morning walk?

4.4.2.2 Fillers

The 40 fillers were also four-sentence narrative discourses. Half of them were fully coherent, and half were not, based on my own judgments. Among the latter half, the decoherence point, i.e., the expected rejection point, varied between sentences 2-4.

The reason for the decoherence varied as well: half involved causal relations, but were marked with the wrong temporal adverbial (e.g., ‘earlier’ in a Result); and half included implausibly long story-time distances. The fillers can be found in Appendix D.3.

4.4.3 Participants

All participants provided informed consent, and were asked to report their age and their native language(s). A total of 49 native English speakers were recruited via Prolific. Participants received $7 USD for completing the study. Twenty-three participants were excluded from analysis for low accuracy on the easy comprehension questions ($<$70%). The remaining 26 participants are included in the analysis.

4.4.4 Procedure

The experiment was presented through Ibex. The stimuli were presented sentence-by-sentence. Participants indicated whether each sentence was a coherent continuation of the discourse by pressing the ‘e’ key (‘Yes’) or the ‘i’ key (‘No’). There was no time limit for this response. If the ‘e’ key was pressed, the trial advanced to the next sentence. If a participant made it through all four sentences of a discourse, they then had to respond to a polar comprehension question, using the same keys (‘e’ for ‘yes’ and ‘i’ for ‘no’). If a participant pressed the ‘i’/‘No’ key in response to any sentence of a discourse, the trial advanced straight to the comprehension question. The comprehension question timed out after being displayed for 7000ms with no response. If participants responded incorrectly or timed out, they received the message, 'Wrong!' No feedback was given for correct responses.

Order of presentation was randomized for each participant in a Latin square design. At the beginning of the experimental session, participants were given written instructions for the task, which included examples and explanations of discourse continuations that were coherent, and others that were incoherent for a variety of reasons. They then completed four practice trials. The experiment typically lasted 30-40 minutes.

4.4.5 Results

The data were analyzed in R, and modeled in Bayesian mixed effects linear regression models using the brms package (Bürkner, 2017; Carpenter et al., 2017). The default priors were used. Models included random slopes and intercepts by participants and items as group-level effects.

4.4.5.1 Rejection rates

The mean rejection rates for the experimental stimuli are plotted in Figure 4.1. Across the board, the rejection rates are quite low. This was not limited to the experimental stimuli, though, as we can see in Figure 4.2, which shows the mean rejection rates for the fillers. We can also see from Figure 4.2 that the rejection rates for *in*coherent fillers were higher than for coherent fillers at the expected positions. This suggests that participants generally understood the task and performed accordingly, despite being overall fairly accommodating.

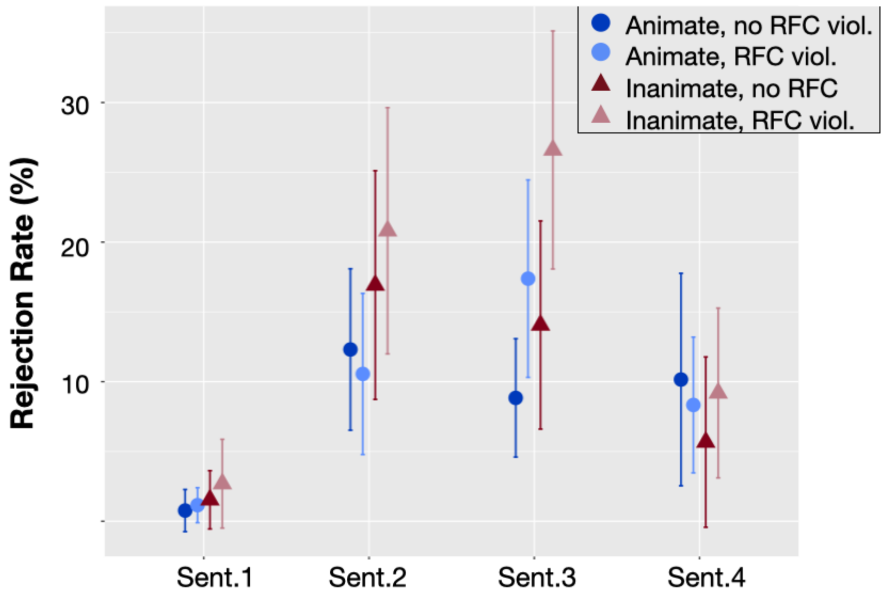


Figure 4.1: Mean rejection rates for experimental stimuli in Experiment 3A. Error bars show a 95% confidence interval.

Fixed effects estimates are given in Table 4.1. At Sentence 2, there is a reliable effect of animacy in the RFC-violation (pronoun) conditions, such that the rejection rate was higher in the Inanimate condition than the Animate. At Sentence 3, there was a reliable RFC effect—RFC-violating conditions were rejected at a higher rate than their counterparts. The same animacy effect from Sentence 2 also obtained here. At Sentence 4, there were no reliable effects, but one—an RFC effect in the Inanimate conditions—came very close.

Table 4.1: Rejection rate analysis: Summary of Bayesian estimates for separate models in Experiment 3A.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Sent. 2	Animate, Definite	Intercept	2.88	0.49	2.00	3.87
		Inanimate	-0.42	0.40	-1.18	0.40
		Pronoun	0.31	0.39	-0.42	1.10
		Inan x Pro	-0.64	0.49	-1.63	0.32
	Inanimate, Pronoun	Intercept	2.01	0.44	1.18	2.94
		Animate	1.31	0.51	0.51	2.23
		Definite	0.43	-0.20	-0.20	1.11
		Anim x Def	-0.60	0.54	-1.69	0.44
Sent. 3	Animate, Definite	Intercept	3.01	0.44	2.23	3.95
		Inanimate	-0.43	0.47	-1.30	0.57
		Pronoun	-1.04	0.42	-1.89	-0.22
		Inan x Pro	-0.26	0.56	-1.39	0.81
	Inanimate, Pronoun	Intercept	1.16	0.29	0.61	1.75
		Animate	0.80	0.34	0.19	1.50
		Definite	1.54	0.50	0.66	2.66
		Anim x Def	0.40	0.76	-0.98	2.04
Sent. 4	Animate, Definite	Intercept	4.42	0.95	2.84	6.55
		Inanimate	1.44	1.34	-0.71	4.54
		Pronoun	0.14	0.80	-1.39	1.82
		Inan x Pro	-0.34	1.53	-3.05	2.98
	Inanimate, Pronoun	Intercept	3.52	0.81	2.21	5.43
		Animate	0.94	0.99	-0.82	3.13
		Definite	1.91	1.20	-0.01	4.79
		Anim x Def	0.41	1.75	-2.50	4.48

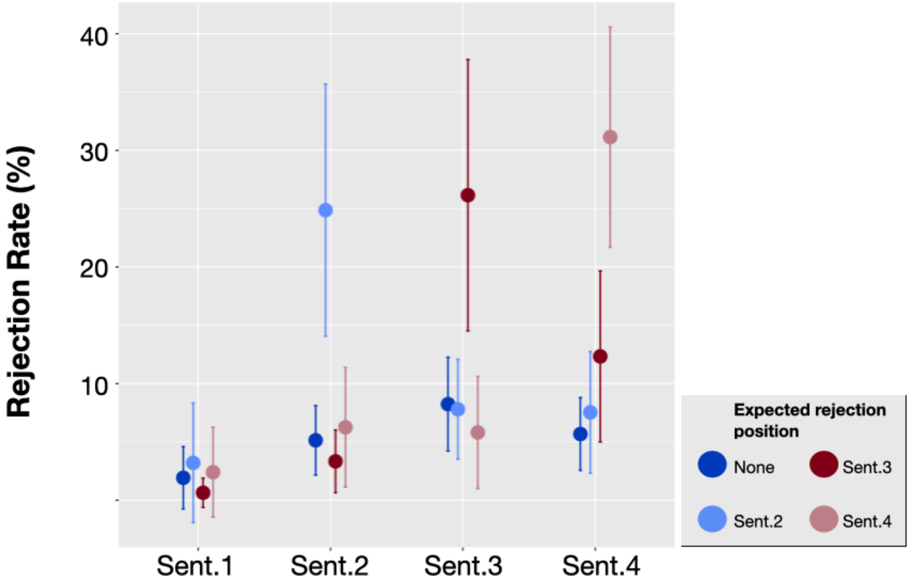


Figure 4.2: Mean rejection rates for fillers in Experiment 3A. Error bars show a 95% confidence interval.

4.4.5.2 Response times

In part because the rejection rates were so low, I also analyzed the response times for those sentences in the experimental stimuli that were accepted. Here, we might hope to see parallels of the effects in the rejection rate data. For instance, we would look for an RFC effect in Sentence 3 such that the response times for the RFC-violating conditions are higher than for the RFC-observing conditions. The mean response times for these sentences are plotted in Figure 4.3.

Our hopes did not come to be, in this case. All of the 95% credible intervals overlapped zero—there were no reliable effects in the response times. The fixed effects estimates are given in Table 4.2. No models were fitted for the rejected sentences be-

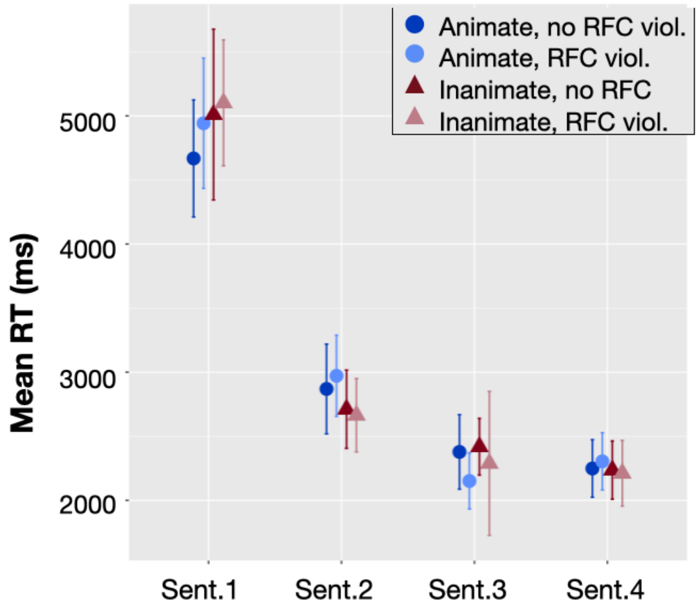


Figure 4.3: Mean rejection rates for accepted experimental sentences, raw, in Experiment 3A. Error bars show a 95% confidence interval.

cause there were too few observations. However, a descriptive comparison of the mean response times for accepted vs. rejected sentences shows that response times were higher for rejected sentences. This pattern obtained for both items (Figure 4.4) and fillers (Figure 4.5).

4.4.6 Discussion

The results of Experiment 3A were, broadly, in line with the original predictions. At Sentence 3—the point at which the RFC factor came into play, we did indeed observe an RFC effect. When Sentence 3 contained a pronoun, whose antecedent was in Sentence 1, which was not on the Right Frontier, the rejection rate was higher than when Sentence

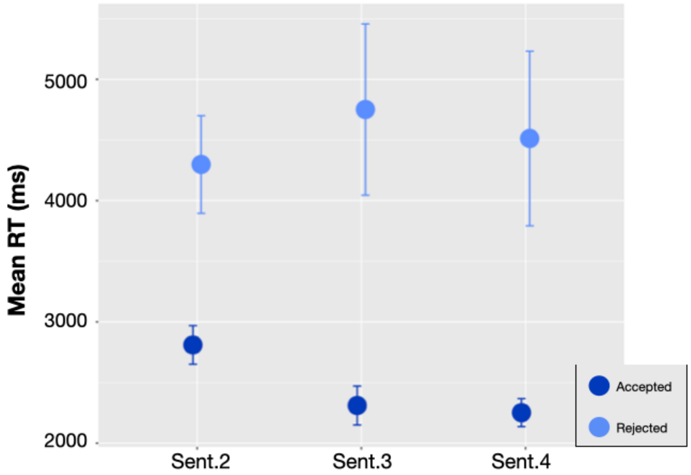


Figure 4.4: Mean RTs for accepted and rejected sentences in items, in Experiment 3A. Error bars show a 95% confidence interval.

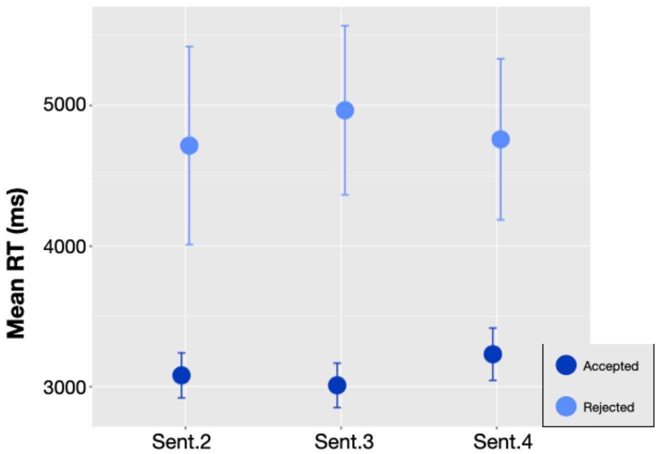


Figure 4.5: Mean RTs for accepted and rejected sentences in fillers, in Experiment 3A. Error bars show a 95% confidence interval.

Table 4.2: Response times analysis: Summary of Bayesian estimates for separate models in Experiment 3A.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Sent. 2	Animate, Definite	Intercept	2899.73	413.22	2083.34	3702.45
		Inanimate	-125.80	174.46	-475.28	211.62
		Pronoun	39.85	166.82	-290.00	373.32
		Inan x Pro	-123.18	223.76	-550.65	323.51
	Inanimate, Pronoun	Intercept	2712.46	371.60	1991.54	3448.69
		Animate	246.80	178.48	-90.32	593.38
		Definite	79.18	180.21	-282.71	433.54
		Anim x Def	-116.97	264.72	-624.89	420.42
Sent. 3	Animate, Definite	Intercept	2533.22	364.28	1830.91	3238.18
		Inanimate	-22.56	270.10	-558.89	514.11
		Pronoun	-171.36	195.24	-553.12	210.93
		Inan x Pro	125.04	419.14	-728.25	935.47
	Inanimate, Pronoun	Intercept	2485.05	335.26	1832.64	3141.15
		Animate	-166.02	344.84	-824.75	511.69
		Definite	13.85	288.56	-537.79	584.42
		Anim x Def	205.02	361.52	-502.81	927.70
Sent. 4	Animate, Definite	Intercept	2474.11	321.11	1839.39	3115.58
		Inanimate	-3.44	200.12	-401.70	394.60
		Pronoun	-6.12	197.82	-436.46	456.37
		Inan x Pro	-95.84	205.11	-484.04	312.54
	Inanimate, Pronoun	Intercept	2379.84	210.84	1965.25	2797.52
		Animate	46.27	195.00	-333.64	436.29
		Definite	59.47	163.55	-253.18	395.44
		Anim x Def	20.98	287.96	-538.31	604.30

3 contained a definite NP instead. In addition to the expected RFC effect, we also observed the predicted Animacy effect at this position: the RFC-violating pronouns with human antecedents (40-b) were rejected at a lower rate than those with inanimate antecedents (40-d). This is compatible with the independent expectation, rooted in the animacy hierarchy literature, that humans are more accessible pronominal referents than inanimates. Further, it is compatible with the way I propose we should view RFC effects, i.e., the combined effects of various factors that modulate salience, rather than a more black-and-white grammatical constraint.

(40) a. **Animate, RFC-observing**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The man** fell over. Jessie circled back to investigate what had happened.

b. **Animate, RFC-violating**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **He** fell over. Jessie circled back to investigate what had happened.

c. **Inanimate, RFC-observing**: Jessie saw **a rickety mailbox** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The mailbox** fell over. Jessie circled back to investigate what had happened.

d. **Inanimate, RFC-violating**: Jessie saw **a rickety mailbox** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **It** fell over. Jessie circled back to investigate what had happened.

Comprehension question (easy): Was Jessie on a morning walk?

Finally, let's not forget that the Animacy effect actually first emerged at Sentence 2. The rejection rate was lower for (41-a) than for (41-b).[6]

(41) a. Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane.

b. Jessie saw **a rickety mailbox** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane.

This effect can also reasonably be categorized as an animacy hierarchy effect, though in a different way from the animacy effect at Sentence 3. The animacy hierarchy also tracks the likelihood that a given entity can be construed as having agency—the higher on the hierarchy, the better as a candidate for agenthood. This could lead to different inferences and expectations about the most coherent possible continuation(s) of a discourse, and thus to the effect we see here. In (41-a), *an elderly man* can be readily inferred to be an independent agent in the world of the narrative. It would then be reasonable for a reader not to form a particularly strong expectation that the man continue

[6]It's not clear why this effect was not reliable for the RFC-observing conditions also, since at this point in the stimuli the Animate conditions were still identical to one another, as were the Inanimate conditions. As we can see in the plot above, though, there is at least a trend in the right direction.

to be significant to the narrative—he is doing his own thing, and Jessie will continue to do hers. In contrast, in (41-b), *a rickety mailbox* cannot be readily inferred to be an agent.[7] Thus, introducing it into the narrative may give rise to an inference that it will figure into the immediate continuation of the narrative—otherwise, it would be irrelevant to have mentioned it at all. So, when the narrative continues—and not only continues to focus on Jessie, but drops the mailbox entirely—it seems incoherent.

4.5 Experiment 3B: Event structure and the RFC

4.5.1 Design and Predictions

This experiment investigated the effects of event structure on long-distance pronominal anaphora resolution. Just as animacy seems to play a role in the salience of an entity, lexical event structure may also modulate entity salience. I hypothesize that the event structure of a transitive verb influences the salience of the entity that fills its Patient/Theme role as follows: the salience of such an entity is based on the comprehenders' world knowledge-based inference of how likely that entity is to come up again in the discourse, depending on the nature of the verb it was originally an argument of.

For instance, consider the two versions of the discourse in (42). Neither of these contains an RFC violation—they are meant to illustrate the hypothesized link between

[7] In this particular item, the meaning of *rickety* may also be influencing matters, because it can give rise to an inference that the rickety thing will collapse imminently, so the narrative ought to continue focusing on that thing. Not all of the Inanimate stimuli were like this, though.

the verb that originally introduces an entity into a discourse and the salience of that entity. The relevant verb and entity are in bold.

(42) a. Jessie **saw an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. The man fell over. Jessie circled back to investigate what had happened.

b. Jessie **brushed by an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. The man fell over. Jessie circled back to investigate what had happened.

For (42-a), I hypothesize that comprehenders will infer lower salience for the entity described by 'an elderly man' when it is the argument of 'saw' than when it is the argument of 'brushed by.' This is because of the distinction between 'saw'—a basic perception verb—and 'brushed by'—which implies closer physical proximity than 'saw,' and therefore a greater potential for continued interaction. Assuming that an entity's salience corresponds directly to its accessibility as a pronoun antecedent, this leads us to the following predictions for the stops-making-sense task: RFC violations should be rejected at a higher rate than non-violations. Further, RFC violations for which the antecedent is the object of 'saw' (or any basic verb of sensing) should be rejected at a higher rate than those for which the antecedent is the object of 'brushed by' (or any verb of contact). This experiment has a 2x2 design, crossing Verb Type {Perception (*saw*), Contact (*brushed by*)} with RFC {RFC-observing, RFC-violating}.

4.5.2 Materials

This experiment included 40 experimental stimuli and 40 fillers. All stimuli were four-sentence narrative discourses. A sample item is shown in (43). The relevant pronouns and their intended antecedents are in bold. The complete materials are provided in Appendix D.2. Per the predictions outlined in the previous section, I predict that (43-a) and (43-c) will not be rejected at all, since they have no RFC violations and are otherwise designed to be coherent throughout. The condition illustrated in (43-b)—an RFC violation in which the antecedent is introduced by a perception verb ('saw'), is predicted to be rejected at the sentence containing the pronoun (i.e., the third sentence) at a nonzero rate. I predict that the (43-d) condition—an RFC violation in which the antecedent is introduced by a verb of contact ('brushed by')—should be rejected at the third sentence at a lower rate than (43-b).

(43) a. **Perception, RFC-observing**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The man** fell over. Jessie circled back to investigate what had happened.

b. **Perception, RFC-violating**: Jessie saw **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **He** fell over. Jessie circled back to investigate what had happened.

c. **Contact, RFC-observing**: Jessie brushed by **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The man** fell over. Jessie circled back to investigate what had happened.

d. **Contact, RFC-violating**: Jessie brushed by **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **He** fell over. Jessie circled back to investigate what had happened.

The fillers were the same as those used in Experiment 3A, and can be found in Appendix D.3.

4.5.3 Participants

All participants provided informed consent, and were asked to report their age and their native language(s). A total of 49 native English speakers were recruited via Prolific. Participants received $7 USD for completing the study. Nineteen participants were excluded from analysis for low accuracy on the easy comprehension questions (<70%). The remaining 30 participants are included in the analysis.

4.5.4 Procedure

The procedure was the same as that of Experiment 3A, which is described in §4.4.4.

4.5.5 Results

The data were analyzed in the same manner as the Experiment 3A data (see §4.4.5).

4.5.5.1 Rejection rates

The mean rejection rates for the experimental stimuli are plotted in Figure 4.6. The mean rejection rates for the fillers are in Figure 4.7. As in Experiment 3A, the rejection rates were rather low overall for both items and fillers. And, also as in Experiment 3A, the rejection rates at incoherent points in fillers were higher than coherent points.

The only reliable effect emerged in Sentence 3: there was an RFC effect for the Contact conditions such that the RFC-violating condition was rejected at a higher rate than the RFC-observing condition. There was a trend towards the same effect for the Perception conditions. Fixed effects estimates are given in Table 4.3.

4.5.5.2 Response times

The mean response times for sentences in experimental stimuli that were accepted are plotted in Figure 4.8.

One reliable effect obtained here: at Sentence 3, there was an RFC effect in the

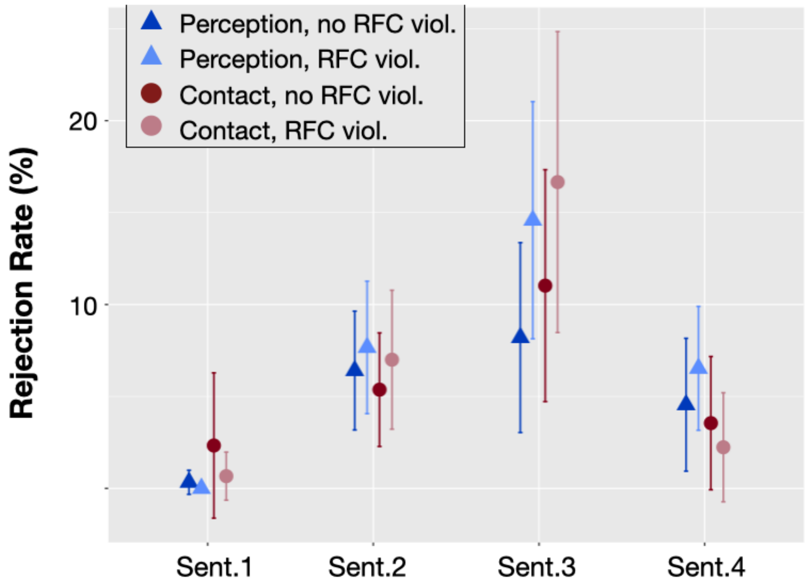


Figure 4.6: Mean rejection rates for experimental stimuli in Experiment 3B. Error bars show a 95% confidence interval.

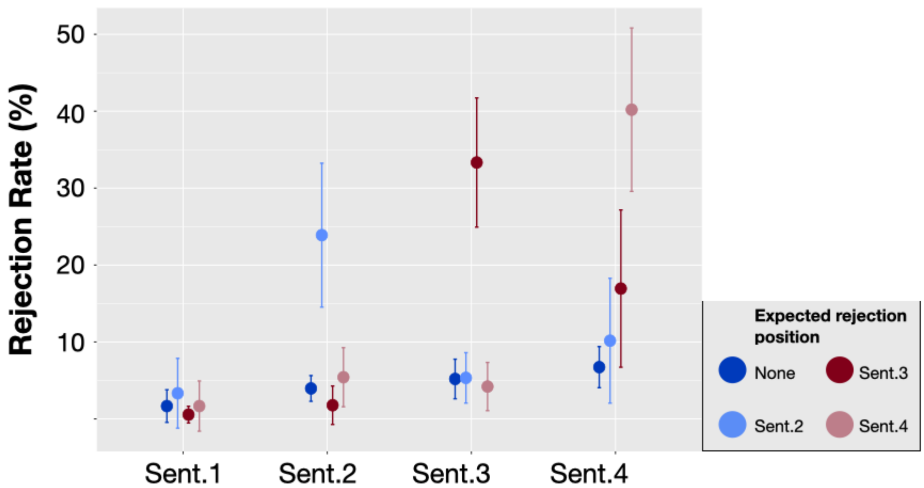


Figure 4.7: Mean rejection rates for fillers in Experiment 3B. Error bars show a 95% confidence interval.

Table 4.3: Rejection rates analysis: Summary of Bayesian estimates for separate models in Experiment 3B.

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Sent. 2	Perception, Definite	Intercept	3.52	0.48	2.65	4.55
		Contact	0.43	0.54	-0.57	1.58
		Pronoun	0.05	0.51	-0.88	1.17
		Contact x Pro	0.27	0.77	-1.14	1.93
	Contact, Pronoun	Intercept	3.61	0.50	2.72	4.69
		Perception	-0.05	0.47	-0.97	0.90
		Definite	0.34	0.47	-0.60	1.29
		Perc x Def	0.13	0.70	-1.14	1.59
Sent. 3	Perception, Definite	Intercept	3.34	0.47	2.52	4.36
		Contact	0.05	0.52	-0.89	1.15
		Pronoun	-0.73	0.42	-1.57	0.09
		Contact x Pro	0.41	0.67	-0.79	1.82
	Contact, Pronoun	Intercept	2.55	0.48	1.69	3.58
		Perception	0.12	0.39	-0.67	0.87
		Definite	0.86	0.42	0.08	1.72
		Perc x Def	0.22	0.59	-0.89	1.46
Sent. 4	Perception, Definite	Intercept	4.17	0.63	3.14	5.59
		Contact	1.42	1.20	-0.46	4.33
		Pronoun	-0.33	0.83	-1.81	1.47
		Contact x Pro	3.84	3.33	-0.28	12.10
	Contact, Pronoun	Intercept	5.06	0.85	3.63	6.96
		Perception	-1.30	0.73	-2.83	0.05
		Definite	1.13	1.36	-1.15	4.19
		Perc x Def	1.01	1.28	-1.40	3.65

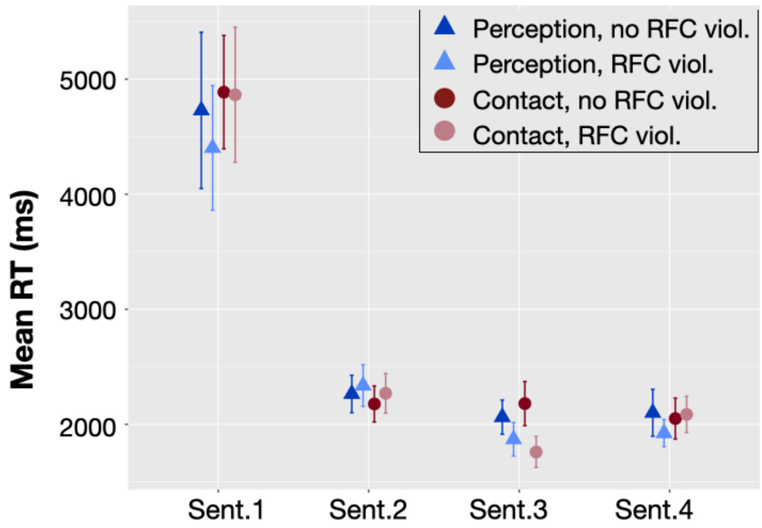


Figure 4.8: Mean response times for accepted experimental sentences, raw, in Experiment 3B. Error bars show a 95% confidence interval.

Contact conditions, but in the opposite of the expected direction. The response times for the RFC-observing condition were higher than for the RFC-violating condition. Fixed effects estimates are given in Table 4.4. No models were fitted for the rejected sentences because there were too few observations. However, as in Experiment 3A, the mean response times were higher for rejected sentences than for accepted sentences for both items (Figure 4.9) and fillers (Figure 4.10).

4.5.6 Discussion

The results of Experiment 3B were somewhat in line with our initial predictions. At Sentence 3, we observed the expected RFC effect, but it was only reliable for the Contact conditions, and was only trending for the Perception conditions. Contrary to

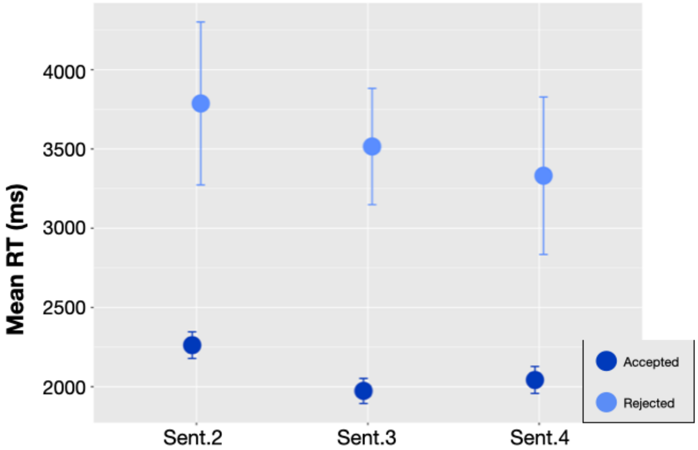


Figure 4.9: Mean RTs for accepted and rejected sentences in items in Experiment 3B. Error bars show a 95% confidence interval.

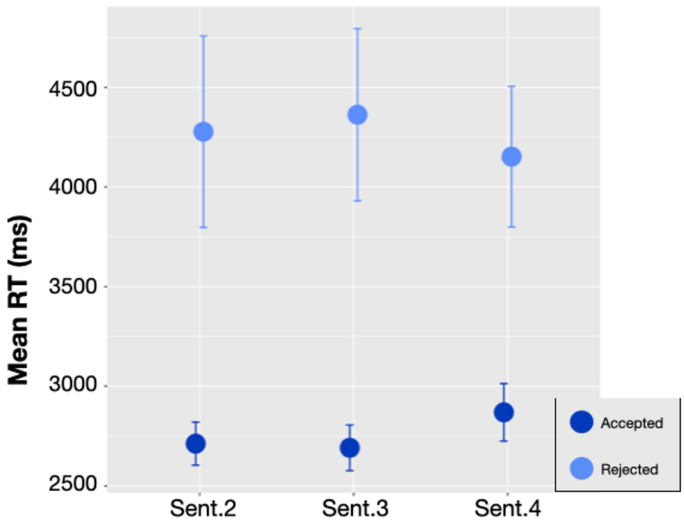


Figure 4.10: Mean RTs for accepted and rejected sentences in fillers in Experiment 3B. Error bars show a 95% confidence interval.

Table 4.4: Response times analysis: Summary of Bayesian estimates for separate models in Experiment 3B

	Reference condition		Estimate	Estimated error	2.5%	97.5%
Sent. 2	Perception, Definite	Intercept	2281.61	156.67	1970.66	2592.08
		Contact	-111.90	116.48	-339.28	115.19
		Pronoun	70.07	116.06	-160.71	298.15
		Contact x Pro	47.35	160.94	-268.73	364.83
	Contact, Pronoun	Intercept	2285.12	157.47	1983.70	2604.47
		Perception	64.06	111.04	-155.27	281.37
		Definite	-122.06	109.13	-331.42	92.73
		Perc x Def	52.47	158.17	-260.92	361.08
Sent. 3	Perception, Definite	Intercept	2102.07	155.79	1790.87	2405.61
		Contact	84.22	105.39	-126.70	285.97
		Pronoun	-161.91	103.28	-366.82	40.89
		Contact x Pro	-226.30	139.28	-495.79	47.17
	Contact, Pronoun	Intercept	1800.41	143.05	1525.23	2079.41
		Perception	142.34	96.86	-50.39	333.95
		Definite	386.87	116.18	149.93	612.82
		Perc x Def	-224.76	141.49	-498.67	61.00
Sent. 4	Perception, Definite	Intercept	2124.04	182.31	1761.42	2481.06
		Contact	-68.11	121.32	-310.40	172.86
		Pronoun	-141.99	110.20	-360.51	74.26
		Contact x Pro	143.35	149.51	-152.87	431.38
	Contact, Pronoun	Intercept	2053.25	152.23	1746.52	2347.13
		Perception	-91.55	105.33	-294.59	116.12
		Definite	-5.26	108.99	-212.33	209.44
		Perc x Def	163.90	176.09	-181.24	521.09

predictions, there was no effect of Verb Type at Sentence 3—we had expected lower rejection rates for Contact RFC violations than for Perception RFC violations, but found no difference either way. Also at Sentence 3, we observed a limited RFC effect in the response times for accepted sentences, such that response times for Contact RFC-observing sentences (44-a) were *slower* than for Contact RFC-violating sentences (44-b). This is the opposite of the expected direction for RFC effects.

(44) a. **Contact, RFC-observing**: Jessie brushed by **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **The man** fell over.

b. **Contact, RFC-violating**: Jessie brushed by **an elderly man** in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. **He** fell over.

The length difference between *the man* and *he* may have contributed to the response time difference, but if this is what happened, it is unclear why the effect didn't also emerge for the Perception conditions, or in the Experiment 3A response times.

There is an interpretation of this result that supports the view that a verb of contact like *brush by* boosts the salience of its Patient/Theme relative to a verb of perception like *see*. First, we must remember that the response times we've analyzed are only those for sentences that were accepted. At Sentence 3, this means that we're dealing with the response times of participants that may not be very sensitive (at least in final-state judgments) to pronominal RFC violations—either at all or just in the trials they accepted. In this context, a definite NP and a pronoun are still in competition with one another, but not in the RFC-observing vs. RFC-violating arena. Instead, we could be seeing something like the repeated name penalty (Gordon, Grosz, & Gilliom, 1993), an effect in which the second sentence in (45-a) is comprehended more slowly than in (45-b).

(45) a. *Atsuko* was hard at work developing a comedy routine. *Atsuko* told Nora a bunch of jokes-in-progress.

b. *Atsuko* was hard at work developing a comedy routine. *She* told Nora a bunch of jokes-in-progress.

In the current experiment, it could be the case that, when introduced by *brushed by*, the *elderly man* was indeed interpreted as highly salient. When this entity is mentioned again in Sentence 3, using *the man* seems redundant. The discourse would still make sense, but may have seemed odd enough that a comprehender would be slower to accept it than a version with a pronoun. Additional analysis and future experiments will be needed to test this, but it is a viable explanation for the current results, and a promising one in that it does point to an event-structural effect on salience in discourse.

4.6 General Discussion and Summary

In this chapter, I focused on operationalizing and testing Asher & Lascarides' SDRT framework. I have argued that the Right Frontier Constraint, which is a structural constraint on discourse-level attachments and pronominal anaphora resolution, provides some of SDRT's clearest, most straightforwardly-testable predictions. In Experiments 3A and 3B, I tested some of those predictions in the domain of pronominal anaphora resolution. I further proposed that what we might simply label the 'RFC effect' from a purely SDRT perspective can actually be decomposed into the effects of various independent factors that influence salience. This raises the possibility that the RFC, which is essentially a grammatical constraint, may not need to be encoded in SDRT, at least with respect to pronominal anaphora in discourse.

In both experiments, we observed the expected RFC effects, in the form of higher rejection rates for sentences that contained RFC-violating pronouns than those that did not. Promisingly, the results of Experiment 3A are compatible with my proposal that 'RFC effects' can be parceled out into SDRT-independent effects. In particular, it seems that the animacy hierarchy plays a role in modulating discourse-level pronominal anaphora resolution. The results of Experiment 3B were not as supportive of my proposal—we observed no interaction between event structure/verb type and RFC effects. It is worth noting, however, that this manipulation was more speculative than the animacy manipulation. The Experiment 3B findings do not rule out the possibility that event structural effects can subsume some RFC effects; they simply suggest that there is not much of a distinction between *see* and *brush by* in this respect. Further, in the apparent reversal of RFC effects that emerged in this experiment, we could still see the possibility that the event structure manipulation did affect salience, but with an outcome that more resembles the repeated name penalty (Gordon et al., 1993) than an RFC violation penalty. Future work will dig more deeply into event structural effects on pronominal anaphora.

Besides the RFC effects, the two experiments had two other results in common. The first was that the rejection rates were very low overall, across items and fillers. This could be interpreted in a few ways. It may be taken to suggest that none of the causes of incoherence in the stimuli, including the RFC violations, are strong 'grammatical' constraints, in the way that some syntactic constraints seem to be. However, this pattern could just as well be taken as an indication that the sentence-by-sentence Stops-Making-Sense paradigm is too coarse-grained to register the full extent of RFC-type effects. In the Maze task or eye-tracking while reading, for instance, we might be better able to observe disruptions in on-line comprehension right at an RFC-violating pronoun. The general lack of response time effects in the current experiments may challenge this view somewhat, but for now the issue remains unresolved.

Considering response times brings us to the other common result across the experiments: the response times for rejected sentences was higher than for accepted sentences. Because these are response times for sentences, we can't tell whether this pattern comes from a disruption in incremental comprehension—e.g., precisely at an RFC-violating pronoun—or from some later integration stage. Further research with more fine-grained methods will have to determine what is happening here. Overall, though, it seems that the Stops-Making-Sense paradigm can be scaled up to the discourse level, and can serve as a useful way to map the search space for more resource-intensive methodologies.

Returning to the status of the RFC, I have tried to be careful in the preceding discussion to separate the notions of the RFC as a constraint on pronominal anaphora and the RFC as a constraint on discourse attachment. The current experiments only dealt with the former usage of the RFC. This was mainly due to pragmatic considerations (the real-life kind, not the linguistic kind). Pronominal anaphora and constraints on it, at the sentence and discourse levels, have been studied previously, so I had some existing work to build on, e.g., in the animacy domain. It is also easier to design well-controlled studies in the RFC-anaphora realm than in the RFC-attachment realm. Future work will tackle the RFC in the discourse-attachment domain.

Chapter 5

Conclusion

This dissertation represents the first steps towards a theory of discourse coherence that integrates the best of two worlds: formal linguistic theories of discourse and structure, and psychological theories of discourse comprehension and its representation in memory. To this end, I have focused on three main questions. These are repeated in (1)-(3).

(1) How does temporal iconicity influence discourse comprehension and discourse recall?

(2) How does causal linking influence discourse comprehension and recall?

(3) What factors govern the comprehension of long-distance relationships, e.g., pronominal anaphora, in discourse?

5.1 Temporal iconicity and 'defaults' in discourse

Tackling the question of temporal iconicity's role in discourse processing required me to consider the notion of 'defaults' in discourse. In both linguistic and psychological discourse theories, represented in this dissertation by Segmented Discourse Representation Theory and Situation Model Theory, Narration—the temporally iconic, non-causal relation—is assumed to be the default relation in discourse. This is a pretty reasonable assumption—it aligns with our lived experience of time, and is free of the added complexity of a causal inference. In a corpus investigation (§2.3), I found that progression does seem to be the most common temporal relation in naturally-occurring written discourse, while backshift—the anti-iconic temporal relation—is both relatively rare and almost always explicitly marked (e.g., with the perfect *had*) when it does occur. When it comes to discourse comprehension, though, there remained an open question about the default nature of progression: Is it a 'strong', first choice default, which comprehenders actively predict during discourse processing? Or is it more of a 'weak', last resort default, which comprehenders only use if all other relations fail? Based on the self-paced reading experiments and Maze task experiment in Chapter 2, I argue that progression is not a strong default. In one of those experiments (Experiment 1B), there was limited evidence to suggest that the plausibility of a backshift interpretation could influence discourse comprehension. The plausibility factor in that experiment was an amalgamation of various pragmatic and lexical semantic factors, e.g, implicit causality. The relations this factor was probing can perhaps be more clearly conceptualized within a formal framework of coherence. To this end, I used some of the core coherence relations of SDRT: Narration, Result, and Explanation. I also shifted the primary experimental focus from on-line comprehension to retrieval. Examining these three coherence relations signified a shift away from studying temporal relations only, towards studying temporal and causal relations, and the interplay between them.

5.2 A leading role for causality?

In Chapter 3, I examine the roles of temporal and causal information in discourse processing within the Situation Model Theory framework. I import the notion of coherence relations from SDRT, largely as a convenient way of labeling/categorizing the intersentential relationships I deal with in this chapter, since SMT is impoverished on this front. The experiments in this chapter are modeled on those of Zwaan (1996). In a set of probe recognition studies, Zwaan investigated the effects of story-time distance on comprehension and situation model representations. In the process, he weighed two views of temporal continuity: Strong Iconity and the Scenario Model. His findings suggested that any temporal jump in a discourse, i.e., anything larger than 'a moment', spurred the creation of a situation model boundary. That is, he found in favor of the Strong Iconicity view. My experiments are modeled on his, and expanded on them by including not only Narrations, but Results and Explanations as well.

My results were not consistent the Strong Iconicity view, nor with the Scenario Model—I found no effects of story-time distance. Considering that my experiments had more statistical power than both Zwaan's and Rinck & Bower's (whose findings supported the Scenario Model), and that those previous effects were quite small, this casts some doubt on the earlier findings. If future work turns up more evidence that the story-time distance effect is not reliable, it may be necessary to rethink how temporal continuity is determined in SMT, and how prominent the temporal dimension is in situation model construction.

My probe experiments did yield a positive finding: probe task performance was better for both Results and Explanations than for Narrations. This, I argue, suggests that Narration may constitute a discontinuity in the causal dimension. I also recast the causal continuity advantage in terms of encoding strength. This gives us a more modern understanding of the difference between Narration and Result, which is predicted in the original framing of SMT. However, on that more traditional view of SMT, in which backshifts are also tacitly assumed to constitute temporal discontinuities, and in which the temporal and causal dimensions are not proposed to influence situation model construction to different degrees, the difference between Narration and Explanation is unexpected. In the encoding strength version, it can be accounted for. On this view, encoding a causal link means greater representational complexity than no causal link, which can strengthen a memory representation.

5.3 Salience on the Right Frontier

In Chapter 4, I turn my focus to SDRT, arguing that the Right Frontier Constraint, especially in its role as a constraint on pronominal anaphora resolution, provides the clearest avenue for operationalizing SDRT. I question the necessity of the RFC, a structural, 'grammatical' constraint, for pronominal anaphora in discourse. I propose that what we call RFC effects within SDRT may actually be a conglomeration of effects of various independent factors that influence pronominal anaphora, perhaps by modulating salience. Two possible factors—animacy and event structure—were explored in a pair of stops-making-sense studies. The findings of the animacy experiment (Experiment 3A) support the view that animacy—in particular, the animacy hierarchy—can modulate discourse-level pronoun resolution. The findings of the event structure experiment were more mixed—they did not suggest that event structure modulated RFC-type effects. However, there was some evidence to suggest that event structure did affect salience, yielding an effect that resembles the repeated name penalty (Gordon et al., 1993) more than an RFC-type effect.

5.4 Future directions

This dissertation is just the beginning of the journey towards a theory of discourse that integrates the best of two worlds: formal linguistic coherence theory and psychological discourse theory. Linguistic theory provides a well-articulated, linguistically-sensitive framework for dividing up the empirical and representational spaces in which we are working. It also captures the ways in which different categories of information, such as temporal and causal information, may interact with one another. It is designed to predict patterns of coherence and coreference in final-state judgments, not to be a model for discourse processing. Psychological theory provides a framework that generates relatively clear predictions about processing, as well as established methodologies for testing those predictions. It is designed to predict patterns of chunking discourse information for storage in memory. It will take considerably more study to work out how best to combine these theories, but we can now provide a few desiderata for an integrated theory. We do not want a theory of discourse processing that involves making hard-and-fast predictions about discourse structure. While relatively strong expectations may be triggered by certain linguistic information, e.g., an implicit causality verb creating a strong expectation of an upcoming Explanation, these are not indefeasible predictions. We do want a theory of which factors—linguistic and otherwise—influence discourse structure, interpretation, and memory representation, of the relative degrees of influence that those factors have, and when they can exert that influence. This will involve implementing experimental designs in methodological pairs (or even triples)—we need measures of incremental processing, such as eye-tracking while reading, as well as measures of retrieval performance, such as probe tasks. In terms of designs, there are myriad possible paths of inquiry, from further investigation of linguistic cues—such as tense and aspect—to the 'squishier' arenas of, e.g., salience and causality.

Appendix A

Norming study: Plausibility of backshift

A.1 Overview

We ran a norming study to establish baseline levels for the plausibility of a backshift interpretation for specific pairs of events. We used these measures to construct the items for Experiments 1A-1D. To measure plausibility, we asked participants to put the sentences of individual items—each of which comprised two sentences—in chronological order. Responses were given in a rating task format. The goal of this task was to get an estimate of the plausibility of a backshift relationship holding between two events. ”Plausibility” is an umbrella for world-knowledge and lexical-semantic factors, which I do not attempt to peel apart further in the current work. The norming task was meant to estimate backshift plausibility for specific pairs of events, as independently as possible from overt cues to temporal relation such as tense/aspect markers or temporal adverbials. Since the task was meant to gauge the effects of world knowledge and lexical semantics on temporal relation, the items used in it are not identical to those used in the self-paced reading experiments. Items in this task were all in simple present tense, and pronouns were not used.

A.2 Methods

A.2.1 Materials

The materials comprise a fully crossed 2 x 3 design, with the factors Order (Same, Reverse) and Plausibility of backshift interpretation (High, Medium, Low). Plausibility was based on my own judgments: first, an S1 was constructed, then three different S2s were constructed, each intended to have a different plausibility of being in a backshift relationship with S1.

To control for possible effects of linear presentation order, we manipulated the linear order: in Same conditions, the sentences were presented in the linear order in which they were originally constructed, and in which they would be presented in the self-paced reading experiments. In Reverse conditions, they were presented in the opposite order.

Forty-two items were constructed, and no fillers were used. These 42 items were split into two sets; 30 participants saw 21 items, and 30 different participants saw the other 21. A sample item is in ((1)).

(1) a. (Same, High): Lianne pokes Eric. Eric falls asleep in class.
    b. (Reverse, High): Eric falls asleep in class. Lianne pokes Eric.
    c. (Same, Medium): Lianne pokes Eric. Eric pinches Lianne.
    d. (Reverse, Medium): Eric pinches Lianne. Lianne pokes Eric.
    e. (Same, Low): Lianne pokes Eric. Eric jumps out of Lianne's reach.
    f. (Reverse, Low): Eric jumps out of Lianne's reach. Lianne pokes Eric.

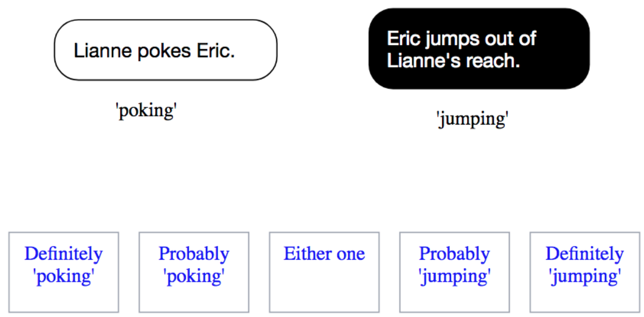
A.2.2 Participants

Sixty participants were recruited through Amazon Mechanical Turk. Participation was restricted to IP addresses in the United States. All participants self-identified as native English speakers and gave informed consent. Participants were paid $2 for completing the task, which took approximately 15 minutes to complete.

A.2.3 Procedure

Each set of 21 items was distributed into six Latin Square lists. The order of presentation was randomized. The experiment was administered online via Ibex. Participants were shown items in a card format, with one sentence per card, as shown in the sample display in ((2)). They were told that the cards were captions for storyboards, and that they had dropped the cards, and now had to put them into chronological order. Participants judged which card should go earlier on the storyboard on a five-point scale.

(2)



At the beginning of the experiment, participants were given written instructions, and completed three practice items.

A.2.4 Analysis and Results

Responses were coded on a 1-5 scale, on which 1 corresponds to Definitely Progression and 5 to Definitely Backshift. The Order manipulation did not have an effect on responses. Means and standard deviations were calculated for each Plausibility level and for each item. Individual stimuli whose means fell outside one standard deviation from the mean of their intended Plausibility level and within one standard deviation of another level were reassigned to that second level for the construction of self-paced reading items. An example of a reassigned stimulus is in ((3)).

(3) Felicia criticizes Gerald. Gerald nags Felicia.

This stimulus was originally generated for the Medium Plausibility condition, but it patterned with the Low Plausibility items, and was thus reassigned to the latter category. A total of 16 stimuli, from 8 different item sets, were reassigned.

Appendix B

Experiment 1 Materials

B.1 Experiment 1A: AdvP Position x Temporal Relation

(1) Eddie fell. |

a. A moment earlier, | on the path | to the lake | behind the house, | Francine tripped | over a loose rock.

b. A moment later, | on the path | to the lake | behind the house, | Francine tripped | over a loose rock.

c. Francine tripped | over a loose rock | a moment earlier, | on the path | to the lake | behind the house.

d. Francine tripped | over a loose rock | a moment later, | on the path | to the lake | behind the house.

(2) Maxine broke her foot. |

a. Prior to that, | at the hospital | down the street | from her apartment, | she got | it x-rayed.

b. Following that, | at the hospital | down the street | from her apartment, | she got | it x-rayed.

c. She got | it x-rayed | prior to that, | at the hospital | down the street | from her apartment.

d. She got | it x-rayed | following that, | at the hospital | down the street | from her apartment.

(3) Felicia criticized Gerald. |

a. The day before that, | in line | at the grocery store | near the bank, | he nagged | her about the chores.

b. The day after that, | in line | at the grocery store | near the bank, | he nagged | her about the chores.

c. He nagged | her about the chores | the day before that, | in line | at the grocery store | near the bank.

d. He nagged | her about the chores | the day after that, | in line | at the grocery store | near the bank.

(4) Charles scolded Howard. |

a. Right before that, | on the path | in the middle | of the park, Howard made | a rude face at him.

b. Right after that, | on the path | in the middle | of the park, Howard made | a rude face at him.

c. Howard made | a rude face at him | right before that, | on the path | in the middle | of the park.

d. Howard made | a rude face at him | right after that, | on the path | in the middle | of the park.

(5) Phoebe yelled at Kevin. |

a. Earlier that day, | at his desk | across the room | in history class, | he shouted | angrily at her.

b. Later that day, | at his desk | across the room | in history class, | he shouted | angrily at her.

c. He shouted | angrily at her | earlier that day, | at his desk | across the room | in history class.

d. He shouted | angrily at her | later that day, | at his desk | across the room | in history class.

(6) Nadya forgave Patrick. |

a. Earlier that day, | in the aisle | of the train | to the city, | he gave | her a big hug.

b. Later that day, | in the aisle | of the train | to the city, | he gave | her a big hug.

c. He gave | her a big hug | earlier that day, | in the aisle | of the train | to the city.

d. He gave | her a big hug | later that day, | in the aisle | of the train | to the city.

(7) Pete greeted Daisy. |

a. Moments before that, | with bright eyes | and an excited grin | on her face, | she told | him the big news.

b. Moments after that, | with bright eyes | and an excited grin | on her face, | she told | him the big news.

c. She told | him the big news | moments before that, | with bright eyes | and an excited grin | on her face.

d. She told | him the big news | moments after that, | with bright eyes | and an excited grin | on her face.

(8) Luke hugged Ruby. |

a. Earlier that afternoon, | at the park | with the dog run | near the beach, | she started | to feel better.

b. Later that afternoon, | at the park | with the dog run | near the beach, | she started | to feel better.

c. She started to feel better | earlier that afternoon, | at the park | with the dog run | near the beach.

d. She started to feel better | later that afternoon, | at the park | with the dog run | near the beach.

(9) Julie voted for Sean. |

a. The previous week, | at the debate | in the auditorium | at the high school, | he promised | to be a good class president.

b. The next week, | at the debate | in the auditorium | at the high school, | he promised | to be a good class president.

He promised | to be a good class president | the previous week, | at the debate | in the auditorium | at the high school.

He promised | to be a good class president | the next week, | at the debate | in the auditorium | at the high school.

(10) Eliza interrupted Bob. |

a. Just before that, | at the conference | at the university | in midtown, | he ignored | her raised hand.

b. Just after that, | at the conference | at the university | in midtown, | he ignored | her raised hand.

c. He ignored | her raised hand | just before that, | at the conference | at the university | in midtown.

d. He ignored | her raised hand | just after that, | at the conference | at the university | in midtown.

(11) Zoey chased the cat. |

a. Immediately before, | with an angry ‘meow’ | and its claws | and fangs bared, | it ran | up a huge tree.

b. Immediately after, | with an angry ‘meow’ | and its claws | and fangs bared, | it ran | up a huge tree.

c. It ran | up a huge tree | immediately before, | with an angry ‘meow’ | and its claws | and fangs bared.

d. It ran | up a huge tree | immediately after, | with an angry ‘meow’ | and its claws | and fangs bared.

(12) Ned forgot his lunch. |

a. An hour before, | in his rush | to the bus stop | two blocks away, | he realized | his shirt was inside out.

b. An hour after, | in his rush | to the bus stop | two blocks away, | he realized | his shirt was inside out.

c. He realized | his shirt was inside out | an hour before, | in his rush | to the bus stop | two blocks away.

d. He realized | his shirt was inside out | an hour after, | in his rush | to the bus stop | two blocks away.

(13) Paul bought a new phone. |

a. The day before, | at the table | by the window | in his kitchen, | he gave | his old phone to his sister.

b. The day after, | at the table | by the window | in his kitchen, | he gave | his old phone to his sister.

c. He gave | his old phone to his sister | the day before, | at the table | by the window | in his kitchen.

d. He gave | his old phone to his sister | the day after, | at the table | by the window | in his kitchen.

(14) Rob consulted Laura. |

a. The previous year, | in the lab | with her team | of graduate students, | she wrote | an influential paper.

b. The next year, | in the lab | with her team | of graduate students, | she wrote | an influential paper.

c. She wrote | an influential paper | the previous year, | in the lab | with her team | of graduate students.

d. She wrote | an influential paper | the next year, | in the lab | with her team | of graduate students.

(15) Mary tricked Edwin. |

a. The previous evening, | in front of | all their friends | at the party, | she discovered | he was quite gullible.

b. The next evening, | in front of | all their friends | at the party, | she discovered | he was quite gullible.

c. She discovered | he was quite gullible | the previous evening, | in front of | all their friends | at the party.

d. She discovered | he was quite gullible | the next evening, | in front of | all their friends | at the party.

(16) Diane divorced Lucy. |

a. Earlier that year, | in the city | by the lake | near her hometown, | Lucy moved | to a new neighborhood.

b. Later that year, | in the city | by the lake | near her hometown, | Lucy moved | to a new neighborhood.

c. Lucy moved | to a new neighborhood | earlier that year, | in the city | by the lake | near her hometown.

d. Lucy moved | to a new neighborhood | later that year, | in the city | by the lake | near her hometown.

(17) Allison hired Jeff. |

a. Prior to that, | with the team | in the offices | on the tenth floor, | he worked on | many important projects.

b. Following that, | with the team | in the offices | on the tenth floor, | he worked on | many important projects.

c. He worked on | many important projects | prior to that, | with the team | in the offices | on the tenth floor.

d. He worked on | many important projects | following that, | with the team | in the offices | on the tenth floor.

(18) Adam consoled Ellen. |

a. A few minutes before, | on the field | at the track meet | in the next town, | she stopped | crying so hard.

b. A few minutes after, | on the field | at the track meet | in the next town, | she stopped | crying so hard.

c. She stopped | crying so hard | a few minutes before, | on the field | at the track meet | in the next town. Ṣhe stopped | crying so hard | a few minutes after, | on the field | at the track meet | in the next town.

(19) Nellie laughed at Oliver. |

a. A moment before, | by the table | in the middle | of the cafeteria, | he stopped | talking to him.

b. A moment after, | by the table | in the middle | of the cafeteria, | he stopped | talking to him.

c. He stopped | talking to her | a moment before, | by the table | in the middle | of the cafeteria.

d. He stopped | talking to her | a moment after, | by the table | in the middle | of the cafeteria.

(20) Emily fixed her laptop. |

a. Earlier that day, | with her discount | at the bookstore | on campus, | she bought | a new laptop sleeve.

b. Later that day, | with her discount | at the bookstore | on campus, | she bought | a new laptop sleeve.

c. She bought | a new laptop sleeve | earlier that day, | with her discount | at the bookstore | on campus.

d. She bought | a new laptop sleeve | later that day, | with her discount | at the bookstore | on campus.

(21) Edith replaced her car tires. |

a. Two days before, | on the highway | in the snow | and the wind, | she drove | safely to work.

b. Two days after, | on the highway | in the snow | and the wind, | she drove | safely to work.

c. She drove | safely to work | two days before, | on the highway | in the snow | and the wind.

d. She drove | safely to work | two days after, | on the highway | in the snow | and the wind.

(22) Joseph found Maya's dog. |

a. Earlier that afternoon, | on the trail | into the woods | outside town, | he saw | her 'Lost Dog' flyers.

b. Later that afternoon, | on the trail | into the woods | outside town, | he saw | her 'Lost Dog' flyers.

c. He saw | her 'Lost Dog' flyers | earlier that afternoon, | on the trail | into the woods | outside town.

d. He saw | her 'Lost Dog' flyers | later that afternoon, | on the trail | into the woods | outside town.

(23) Katie scared the squirrel. |

a. A moment before, | from its perch | in the tree | by the fence, | it squeaked | loudly at her.

b. A moment after, | from its perch | in the tree | by the fence, | it squeaked | loudly at her.

c. It squeaked | loudly at her | a moment before, | from its perch | in the tree | by the fence.

d. It squeaked | loudly at her | a moment after, | from its perch | in the tree | by the fence.

(24) Jimmy cleaned his house. |

a. The previous night, | with great enthusiasm | and no regard | for his dismay, | his kids tracked | mud through the hall.

b. The next night, | with great enthusiasm | and no regard | for his dismay, | his kids tracked | mud through the hall.

c. His kids tracked | mud through the hall | the previous night, | with great enthusiasm | and no regard | for his dismay.

d. His kids tracked | mud through the hall | the next night, | with great enthusiasm | and no regard | for his dismay.

(25) Lianne poked Eric. |

a. A moment earlier, | in the aisle | near the back | of the lecture hall, | he pinched | her on the arm.

b. A moment later, | in the aisle | near the back | of the lecture hall, | he pinched | her on the arm.

c. He pinched | her on the arm | a moment earlier, | in the aisle | near the back | of the lecture hall.

d. He pinched | her on the arm | a moment later, | in the aisle | near the back | of the lecture hall.

(26) Gideon weighed his dinner. |

a. Prior to that, | with the advice | of his doctor | in mind, | he added | more carrots | to his plate.

b. Following that, | with the advice | of his doctor | in mind, | he added | more carrots | to his plate.

c. He added | more carrots | to his plate | prior to that, | with the advice | of his doctor | in mind.

d. He added | more carrots | to his plate | following that, | with the advice | of his doctor | in mind.

(27) Simon called Nicole. |

a. Prior to that, | at a restaurant | in the middle | of the neighborhood, | she asked | him to call her back.

b. Following that, | at a restaurant | in the middle | of the neighborhood, | she asked | him to call her back.

c. She asked | him to call her back | prior to that, | at a restaurant | in the middle | of the neighborhood.

d. She asked | him to call her back | following that, | at a restaurant | in the middle | of the neighborhood.

(28) Andrea complimented Zach. |

a. Just before that, | in the shade | on the patio | in his yard, | he felt | his bad mood disappear.

b. Just after that, | in the shade | on the patio | in his yard, | he felt | his bad mood disappear.

c. He felt | his bad mood disappear | just before that, | in the shade | on the patio | in his yard.

d. He felt | his bad mood disappear | just after that, | in the shade | on the patio | in his yard.

(29) Harry abandoned his car. |

a. An hour earlier, | by the side | of the highway | through the mountains, | he called | a towing service.

b. An hour later, | by the side | of the highway | through the mountains, | he called | a towing service.

c. He called | a towing service | an hour earlier, | by the side | of the highway | through the mountains.

d. He called | a towing service | an hour later, | by the side | of the highway | through the mountains.

(30) Omar changed his clothes. |

a. Earlier that day, | in the exhibition | of modern art | at the museum, | he noticed | a hole in his sweater.

b. Later that day, | in the exhibition | of modern art | at the museum, | he noticed | a hole in his sweater.

c. He noticed | a hole in his sweater | earlier that day, | in the exhibition | of modern art | at the museum.

d. He noticed | a hole in his sweater | later that day, | in the exhibition | of modern art | at the museum.

(31) Millie sneezed. |

a. A few seconds before, | at a busy cafe | full of students | and professors, | she blew | her nose loudly.

b. A few seconds after, | at a busy cafe | full of students | and professors,

| she blew | her nose loudly.

c. She blew | her nose loudly | a few seconds before, | at a busy cafe | full of students | and professors.

d. She blew | her nose loudly | a few seconds after, | at a busy cafe | full of students | and professors.

(32) Celeste quit surfing. |

a. The previous month, | with some friends | from her class | at the gym, | she joined | a softball league.

b. The next month, | with some friends | from her class | at the gym, | she joined | a softball league.

c. She joined | a softball league | the previous month, | with some friends | from her class | at the gym.

d. She joined | a softball league | the next month, | with some friends | from her class | at the gym.

(33) Dwayne rented a car. |

a. The previous morning, | with a couple | of old friends | from college, | he arrived | in Las Vegas.

b. The next morning, | with a couple | of old friends | from college, | he arrived | in Las Vegas.

c. He arrived | in Las Vegas | the previous morning, | with a couple | of old friends | from college.

d. He arrived | in Las Vegas | the next morning, | with a couple | of old friends | from college.

(34) Paige misread a street sign. |

a. Moments earlier, | in her rush | to band rehearsal | at the school, | she took | a wrong turn.

b. Moments later, | in her rush | to band rehearsal | at the school, | she took | a wrong turn.

c. She took | a wrong turn | moments earlier, | in her rush | to band rehearsal | at the school.

d. She took | a wrong turn | moments later, | in her rush | to band rehearsal | at the school.

(35) Luisa scrubbed her kitchen counter. |

a. Before that, | in preparation | for the party | at her house, | she vacuumed | all the carpets.

b. After that, | in preparation | for the party | at her house, | she vacuumed | all the carpets.

c. She vacuumed | all the carpets | before that, | in preparation | for the party

| at her house.

d. She vacuumed | all the carpets | She vacuumed | all the carpets | before that, | in preparation | for the party | at her house | after that, | in preparation | for the party | at her house.

(36) Larry failed the midterm exam. |

a. A few days earlier, | at a desk | in the library | by himself, | he stayed up | all night worrying.

b. A few days later, | at a desk | in the library | by himself, | he stayed up | all night worrying.

c. He stayed up | all night worrying | a few days earlier, | at a desk | in the library | by himself.

d. He stayed up | all night worrying | a few days later, | at a desk | in the library | by himself.

B.2 Experiment 1B: Plausibility x Temporal Relation

(37) Eddie fell. |

a. Francine pushed | him from behind | a moment earlier, | on the path | to the lake | behind the house.

b. Francine pushed | him from behind | a moment later, | on the path | to the lake | behind the house.

c. Francine tripped | over a loose rock | a moment earlier, | on the path | to the lake | behind the house.

d. Francine tripped | over a loose rock | a moment later, | on the path | to the lake | behind the house.

(38) Maxine broke her foot. |

a. She fell | awkwardly on it | prior to that, | at the hospital | down the street | from her apartment.

b. She fell | awkwardly on it | following that, | at the hospital | down the street | from her apartment.

c. She got | it x-rayed | prior to that, | at the hospital | down the street | from her apartment.

d. She got | it x-rayed | following that, | at the hospital | down the street | from her apartment.

(39) Felicia criticized Gerald. |

a. He bought | the wrong bread | the day before that, | in line | at the grocery store | near the bank.

b. He bought | the wrong bread | the day after that, | in line | at the grocery store | near the bank.

c. He nagged | her about the chores | the day before that, | in line | at the grocery store | near the bank.

d. He nagged | her about the chores | the day after that, | in line | at the grocery store | near the bank.

(40) Charles scolded Howard. |

a. Howard made | an enormous mess | right before that, | on the path | in the middle | of the park.

b. Howard made | an enormous mess | right after that, | on the path | in the middle | of the park.

c. Howard walked | away quickly | right before that, | on the path | in the middle | of the park.

d. Howard walked | away quickly | right after that, | on the path | in the middle | of the park.

(41) Phoebe yelled at Kevin. |

a. He stole | her favorite pencil | earlier that day, | from her desk | across the room | in history class.

b. He stole | her favorite pencil | later that day, | from her desk | across the room | in history class.

c. He burst | into hysterical tears | earlier that day, | at his desk | across the room | in history class.

d. He burst | into hysterical tears | later that day, | at his desk | across the room | in history class.

(42) Nadya forgave Patrick. |

a. He apologized | profusely to her | earlier that day, | in the aisle | of the train | to the city.

b. He apologized | profusely to her | later that day, | in the aisle | of the train | to the city.

c. He breathed | a big sigh of relief | earlier that day, | in the aisle | of the train | to the city.

d. He breathed | a big sigh of relief | later that day, | in the aisle | of the train | to the city.

(43) Pete greeted Daisy. |

a. She ran | into the room | moments before that, | with bright eyes | and an excited grin | on her face.

b. She ran | into the room | moments after that, | with bright eyes | and an excited grin | on her face.

c. She waved | happily to him | moments before that, | with bright eyes | and an excited grin | on her face.

d. She waved | happily to him | moments after that, | with bright eyes | and an excited grin | on her face.

(44) Luke hugged Ruby. |

a. She scored | the winning goal | earlier that afternoon, | at the park | with the dog run | near the beach.

b. She scored | the winning goal | later that afternoon, | at the park | with the dog run | near the beach.

c. She tousled | his hair playfully | earlier that afternoon, | at the park | with the dog run | near the beach.

d. She tousled | his hair playfully | later that afternoon, | at the park | with the dog run | near the beach.

(45) Eliza interrupted Bob. |

a. He went over | his speech time | just before that, | at the conference | at the university | in midtown.

b. He went over | his speech time | just after that, | at the conference | at the university | in midtown.

c. He ignored | her raised hand | just before that, | at the conference | at the university | in midtown.

d. He ignored | her raised hand | just after that, | at the conference | at the university | in midtown.

(46) Zoey chased the cat. |

a. It escaped | her apartment | immediately before, | with an angry ‘meow’ | and its claws | and fangs bared.

b. It escaped | her apartment | immediately after, | with an angry ‘meow’ | and its claws | and fangs bared.

c. It ran | up a huge tree | immediately before, | with an angry ‘meow’ | and its claws | and fangs bared.

d. It ran | up a huge tree | immediately after, | with an angry ‘meow’ | and its claws | and fangs bared.

(47) Ned forgot his lunch. |

a. He packed | soup and a sandwich | an hour before, | in his rush | to the bus stop | two blocks away.

b. He packed | soup and a sandwich | an hour after, | in his rush | to the bus stop | two blocks away.

c. He realized | his shirt was on inside out | an hour before, | in his rush | to the bus stop | two blocks away.

d. He realized | his shirt was on inside out | an hour after, | in his rush | to the bus stop | two blocks away.

(48) Paul bought a new phone. |

a. He lost | his old phone with the cracked screen | the day before, | on the trail | in the woods | behind his house.

b. He lost | his old phone with the cracked screen | the day after, | on the trail | in the woods | behind his house.

c. He imported | the contacts from his old phone | the day before, | at the table | by the window | in his kitchen.

d. He imported | the contacts from his old phone | the day after, | at the table | by the window | in his kitchen.

(49) Rob consulted Laura. |

a. She stopped by | his office | the previous morning, | in the suite | down the hall | from her lab.

b. She stopped by | his office | the next morning, | in the suite | down the hall | from her lab.

c. She referred | him to someone else | the previous morning, | in the suite | down the hall | from his lab.

d. She referred | him to someone else | the next morning, | in the suite | down the hall | from his lab.

(50) Mary tricked Edwin. |

a. She made | a very elaborate plan | the previous evening, | in front of | all their friends | at the party.

b. She made | a very elaborate plan | the next evening, | in front of | all their friends | at the party.

c. She discovered | he was quite gullible | the previous evening, | in front of | all their friends | at the party.

d. She discovered | he was quite gullible | the next evening, | in front of | all their friends | at the party.

(51) Diane divorced Lucy. |

a. Lucy cheated | her with an old friend | earlier that year, | in the city | by the lake | near her hometown.

b. Lucy cheated | her with an old friend | later that year, | in the city | by the lake | near her hometown.

c. Lucy got married | to Charlotte | earlier that year, | in the city | by the lake | near her hometown.

d. Lucy got married | to Charlotte | later that year, | in the city | by the lake | near her hometown.

(52) Allison hired Jeff. |

a. He gave | great answers in his interview | prior to that, | with the team | in the offices | on the tenth floor.

b. He gave | great answers in his interview | following that, | with the team | in the offices | on the tenth floor.

c. He got | the office next to hers | prior to that, with the team | in the offices | on the tenth floor.

d. He got | the office next to hers | following that, with the team | in the offices | on the tenth floor.

(53) Adam consoled Ellen. |

a. She lost | the race by inches | a few minutes before, | on the field | at the track meet | in the next town.

b. She lost | the race by inches | a few minutes after, | on the field | at the track meet | in the next town.

c. She began | crying even harder | a few minutes before, | on the field | at the track meet | in the next town.

d. She began | crying even harder | a few minutes after, | on the field | at the track meet | in the next town.

(54) Nellie laughed at Oliver. |

a. He slipped | on a banana peel | a moment before, | by the table | in the middle | of the cafeteria.

b. He slipped | on a banana peel | a moment after, | by the table | in the middle | of the cafeteria.

c. He stopped | talking to her | a moment before, | by the table | in the middle | of the cafeteria.

d. He stopped | talking to her | a moment after, | by the table | in the middle | of the cafeteria.

(55) Emily fixed her laptop. |

a. She took | a computer repair class | earlier that year, | in the computer lab | at the bookstore | on campus.

b. She took | a computer repair class | later that year, | in the computer lab | at the bookstore | on campus.

c. She bought | a new laptop sleeve | earlier that year, | in the computer lab | at the bookstore | on campus.

d. She bought | a new laptop sleeve | later that year, | in the computer lab | at the bookstore | on campus.

(56) Edith replaced her car tires. |

a. She noticed | the treads were worn | two days before, | on the highway | in the snow | and the wind.

b. She noticed | the treads were worn | two days after, | on the highway | in the snow | and the wind.

c. She drove | safely to work | two days before, | on the highway | in the snow | and the wind.

d. She drove | safely to work | two days after, | on the highway | in the snow | and the wind.

(57) Joseph found Maya’s dog. |

a. The dog got | off its leash | earlier that afternoon, | on the trail | into the woods | outside town.

b. The dog got | off its leash | later that afternoon, | on the trail | into the woods | outside town.

c. He saw | her ‘Lost Dog’ flyers | earlier that afternoon, | on the trail | into the woods | outside town.

d. He saw | her ‘Lost Dog’ flyers | later that afternoon, | on the trail | into the woods | outside town.

(58) Katie scared the squirrel. |

a. It snuck | into her garden | a moment before, | from its perch | in the tree | by the fence.

b. It snuck | into her garden | a moment after, | from its perch | in the tree | by the fence.

c. It squeaked | loudly at her | a moment before, | from its perch | in the tree | by the fence.

d. It squeaked | loudly at her | a moment after, | from its perch | in the tree | by the fence.

(59) Jimmy cleaned his house. |

a. His parents said | they would visit him | the previous night, | with great enthusiasm | and no regard | for his dismay.

b. His parents said | they would visit him | the next night, | with great enthusiasm | and no regard | for his dismay.

c. His kids tracked | mud through the hall | the previous night, | with great enthusiasm | and no regard | for his dismay.

d. His kids tracked | mud through the hall | the next night, | with great enthusiasm | and no regard | for his dismay.

(60) Lianne poked Eric. |

a. He fell asleep | in biology class | a moment earlier, | in his seat | near the back | of the lecture hall.

b. He fell asleep | in biology class | a moment later, | in his seat | near the back | of the lecture hall.

c. He pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

d. He pinched | her on the arm | a moment later, | in his seat | near the back | of the lecture hall.

(61) Gideon weighed his dinner. |

a. He started | a new low-fat diet | prior to that, | with the advice | of his doctor | in mind.

b. He started | a new low-fat diet | following that, | with the advice | of his doctor | in mind.

c. He added | more carrots to his plate | prior to that, | with the advice | of his doctor | in mind.

d. He added | more carrots to his plate | following that, | with the advice | of his doctor | in mind.

(62) Simon called Nicole. |

a. She gave | her number to him | prior to that, | at a restaurant | in the middle | of the neighborhood.

b. She gave | her number to him | following that, | at a restaurant | in the middle | of the neighborhood.

c. She asked | him to call her back | prior to that, | at a restaurant | in the middle | of the neighborhood.

d. She asked | him to call her back | following that, | at a restaurant | in the middle | of the neighborhood.

(63) Andrea complimented Zach. |

a. He did | a complicated dance move | just before that, | in the shade | on the patio | in his yard.

b. He did | a complicated dance move | just after that, | in the shade | on the patio | in his yard.

c. He said | he liked her jacket | just before that, | in the shade | on the patio | in his yard.

d. He said | he liked her jacket | just after that, | in the shade | on the patio | in his yard.

(64) Harry abandoned his car. |

a. It got stuck | in a snowbank | an hour earlier, | on the side | of the highway | through the mountains.

b. It got stuck | in a snowbank | an hour later, | on the side | of the highway | through the mountains.

c. He called | a towing service | an hour earlier, | on the side | of the highway | through the mountains.

d. He called | a towing service | an hour later, | on the side | of the highway

| through the mountains.

(65) Caleb fired Denise. |

a. She embezzled | from his company | earlier that week, | at her cubicle | in the middle | of the crowded office.

b. She embezzled | from his company | later that week, | at her cubicle | in the middle | of the crowded office.

c. She cleaned | her messy desk | earlier that week, | at her cubicle | in the middle | of the crowded office.

d. She cleaned | her messy desk | later that week, | at her cubicle | in the middle | of the crowded office.

(66) Omar changed his clothes. |

a. He got splattered | with mud | earlier that day, | near the exhibition | of modern art | at the museum.

b. He got splattered | with mud | later that day, | near the exhibition | of modern art | at the museum.

c. Omar changed his clothes. | He went out | on a date with Liz | earlier that day, | at the exhibition | of modern art | at the museum.

d. Omar changed his clothes. | He went out | on a date with Liz | later that day, | at the exhibition | of modern art | at the museum.

(67) Millie sneezed. |

a. She caught | a nasty cold | right before that, | at a busy cafe | full of students | and professors.

b. She caught | a nasty cold | right after that, | at a busy cafe | full of students | and professors.

c. She started | sniffling miserably | right before that, | at a busy cafe | full of students | and professors.

d. She started | sniffling miserably | right after that, | at a busy cafe | full of students | and professors.

(68) Dwayne rented a car. |

a. He wrecked | his old car | the previous morning, | with a couple | of old friends | from college.

b. He wrecked | his old car | the next morning, | with a couple | of old friends | from college.

c. He arrived | in Las Vegas | the previous morning, | with a couple | of old friends | from college.

d. He arrived | in Las Vegas | the next morning, | with a couple | of old friends | from college.

(69) Paige misread a street sign. |

a. She forgot | her glasses at home | minutes earlier, | in her rush | to band rehearsal | at the school.

b. She forgot | her glasses at home | minutes later, | in her rush | to band rehearsal | at the school.

c. She took | a wrong turn | minutes earlier, | in her rush | to band rehearsal | at the school.

d. She took | a wrong turn | minutes later, | in her rush | to band rehearsal | at the school.

(70) Luisa scrubbed her kitchen counter. |

a. She spilled | some soup | before that, | in an anxious mood | about the party | at her house.

b. She spilled | some soup | after that, | in an anxious mood | about the party | at her house.

c. She vacuumed | all the carpets | before that, | in an anxious mood | about the party | at her house.

d. She vacuumed | all the carpets | after that, | in an anxious mood | about the party | at her house.

B.3 Experiment 1C: Aspect x Temporal Relation

(71) Eddie fell. |

a. Francine tripped | over a loose rock | a moment earlier, | on the path | to the lake | behind the house.

b. Francine tripped | over a loose rock | a moment later, | on the path | to the lake | behind the house.

c. Francine had tripped | over a loose rock | a moment earlier, | on the path | to the lake | behind the house.

(72) Maxine broke her foot. |

a. She got | it x-rayed | prior to that, | at the hospital | down the street | from her apartment.

b. She got | it x-rayed | following that, | at the hospital | down the street | from her apartment.

c. She had gotten | it x-rayed | prior to that, | at the hospital | down the street | from her apartment.

(73) Felicia criticized Gerald. |

a. He nagged | her about the chores | the day before that, | in line | at the grocery store | near the bank.

b. He nagged | her about the chores | the day after that, | in line | at the grocery store | near the bank.

c. He had nagged | her about the chores | the day before that, | in line | at the grocery store | near the bank.

(74) Charles scolded Howard. |

a. Howard made | a rude face at him | right before that, | on the path | in the middle | of the park.

b. Howard made | a rude face at him | right after that, | on the path | in the middle | of the park.

c. Howard had made | a rude face at him | right before that, | on the path | in the middle | of the park.

(75) Phoebe yelled at Kevin. |

a. He shouted | angrily at her | earlier that day, | from his seat | across the room | in history class.

b. He shouted | angrily at her | later that day, | from his seat | across the room | in history class.

c. He had shouted | angrily at her | earlier that day, | from his seat | across the room | in history class.

(76) Nadya forgave Patrick. |

a. He gave | her a big hug | earlier that day, | in the aisle | of the train | to the city.

b. He gave | her a big hug | later that day, | in the aisle | of the train | to the city.

c. He had given | her a big hug | earlier that day, | in the aisle | of the train | to the city.

(77) Pete greeted Daisy. |

a. She told | him the big news | moments before that, | with bright eyes | and an excited grin | on her face.

b. She told | him the big news | moments after that, | with bright eyes | and an excited grin | on her face.

c. She had told | him the big news | moments before that, | with bright eyes | and an excited grin | on her face.

(78) Luke hugged Ruby. |

a. She started | to feel better | earlier that afternoon, | at the park | with the dog run | near the beach.

b. She started | to feel better | later that afternoon, | at the park | with the dog run | near the beach.

c. She had started | to feel better | earlier that afternoon, | at the park | with the dog run | near the beach.

(79) Eliza interrupted Bob. |

a. He ignored | her raised hand | just before that, | at the conference | at the university | in midtown.

b. He ignored | her raised hand | just after that, | at the conference | at the university | in midtown.

c. He had ignored | her raised hand | just before that, | at the conference | at the university | in midtown.

(80) Zoey chased the cat. |

a. It ran | up a huge tree | immediately before, | with an angry ‘meow’ | and its claws | and fangs bared.

b. It ran | up a huge tree | immediately after, | with an angry ‘meow’ | and its claws | and fangs bared.

c. It had run | up a huge tree | immediately before, | with an angry ‘meow’ | and its claws | and fangs bared.

(81) Ned forgot his lunch. |

a. He realized | his shirt was inside out | an hour before, | in his rush | to the bus stop | two blocks away.

b. He realized | his shirt was inside out | an hour after, | in his rush | to the bus stop | two blocks away.

c. He had realized | his shirt was inside out | an hour before, | in his rush | to the bus stop | two blocks away.

(82) Paul bought a new phone. |

a. He gave | his old phone to his sister | the day before, | at the table | by the window | in his kitchen.

b. He gave | his old phone to his sister | the day after, | at the table | by the window | in his kitchen.

c. He had given | his old phone to his sister | the day before, | at the table | by the window | in his kitchen.

(83) Rob consulted Laura. |

a. She wrote | an influential paper | the previous year, | in the lab | with her team | of graduate students.

b. She wrote | an influential paper | the next year, | in the lab | with her team | of graduate students.

c. She had written | an influential paper | the previous year, | in the lab | with her team | of graduate students.

(84) Mary tricked Edwin. |

a. She discovered | he was quite gullible | the previous evening, | in front of | all their friends | at the party.

b. She discovered | he was quite gullible | the next evening, | in front of | all their friends | at the party.

c. She had discovered | he was quite gullible | the previous evening, | in front of | all their friends | at the party.

(85) Diane divorced Lucy. |

a. Lucy moved | to a new neighborhood | earlier that year, | in the city | by the lake | near her hometown.

b. Lucy moved | to a new neighborhood | later that year, | in the city | by the lake | near her hometown.

c. Lucy had moved | to a new neighborhood | earlier that year, | in the city | by the lake | near her hometown.

(86) Allison hired Jeff. |

a. He worked on | many important projects | prior to that, | with the team | in the offices | on the tenth floor.

b. He worked on | many important projects | following that, | with the team | in the offices | on the tenth floor.

c. He had worked on | many important projects | prior to that, | with the team | in the offices | on the tenth floor.

(87) Adam consoled Ellen. |

a. She stopped | crying so hard | a few minutes before, | on the field | at the track meet | in the next town.

b. She stopped | crying so hard | a few minutes after, | on the field | at the track meet | in the next town.

c. She had stopped | crying so hard | a few minutes before, | on the field | at the track meet | in the next town.

(88) Nellie laughed at Oliver. |

a. He stopped | talking to her | a moment before, | by the table | in the middle | of the cafeteria.

b. He stopped | talking to her | a moment after, | by the table | in the middle | of the cafeteria.

c. He had stopped | talking to her | a moment before, | by the table | in the middle | of the cafeteria.

(89) Emily fixed her laptop. |

a. She bought | a new laptop sleeve | earlier that year, | in the computer lab | at the bookstore | on campus.

b. She bought | a new laptop sleeve | later that year, | in the computer lab | at the bookstore | on campus.

c. She had bought | a new laptop sleeve | earlier that year, | in the computer lab | at the bookstore | on campus.

(90) Edith replaced her car tires. |

a. She drove | safely to work | two days before, | on the highway | in the snow | and the wind.

b. She drove | safely to work | two days after, | on the highway | in the snow | and the wind.

c. She had driven | safely to work | two days before, | on the highway | in the snow | and the wind.

(91) Joseph found Maya's dog. |

a. He saw | her 'Lost Dog' flyers | earlier that afternoon, | on the trail | into the woods | outside town.

b. He saw | her 'Lost Dog' flyers | later that afternoon, | on the trail | into the woods | outside town.

c. He had seen | her 'Lost Dog' flyers | earlier that afternoon, | on the trail | into the woods | outside town.

(92) Katie scared the squirrel. |

a. It squeaked | loudly at her | a moment before, | from its perch | in the tree | by the fence.

b. It squeaked | loudly at her | a moment after, | from its perch | in the tree

| by the fence.

c. It had squeaked | loudly at her | a moment before, | from its perch | in the tree | by the fence.

(93) Jimmy cleaned his house. |

a. His kids tracked | mud through the hall | the previous night, | with great enthusiasm | and no regard | for his dismay.

b. His kids tracked | mud through the hall | the next night, | with great enthusiasm | and no regard | for his dismay.

c. His kids had tracked | mud through the hall | the previous night, | with great enthusiasm | and no regard | for his dismay.

(94) Lianne poked Eric. |

a. He pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

b. He pinched | her on the arm | a moment later, | in his seat | near the back | of the lecture hall.

c. He had pinched | her on the arm | a moment earlier, | in his seat | near the back | of the lecture hall.

(95) Gideon weighed his dinner. |

a. He added | more carrots to his plate | prior to that, | with the advice | of his doctor | in mind.

b. He added | more carrots to his plate | following that, | with the advice | of his doctor | in mind.

c. He had added | more carrots to his plate | prior to that, | with the advice | of his doctor | in mind.

(96) Simon called Nicole. |

a. She asked | him to call her back | prior to that, | at a restaurant | in the middle | of the neighborhood.

b. She asked | him to call her back | following that, | at a restaurant | in the middle | of the neighborhood.

c. She had asked | him to call her back | prior to that, | at a restaurant | in the middle | of the neighborhood.

(97) Harry abandoned his car. |

a. He called | a towing service | an hour earlier, | on the side | of the highway | through the mountains.

b. He called | a towing service | an hour later, | on the side | of the highway | through the mountains.

c. He had called | a towing service | an hour earlier, | on the side | of the highway | through the mountains.

(98) Omar changed his clothes. |

a. He noticed | a hole in his sweater | earlier that day, | in the exhibition | of modern art | at the museum.

b. He noticed | a hole in his sweater | later that day, | in the exhibition | of modern art | at the museum.

c. He had noticed | a hole in his sweater | earlier that day, | in the exhibition | of modern art | at the museum.

(99) Millie sneezed. |

a. She blew | her nose loudly | right before that, | at a busy cafe | full of students | and professors.

b. She blew | her nose loudly | right before that, | at a busy cafe | full of students | and professors.

c. She had blown | her nose loudly | right before that, | at a busy cafe | full of students | and professors.

(100) Celeste quit surfing. |

a. She joined | a softball league | the previous month, | with some friends | from her class | at the gym.

b. She joined | a softball league | the next month, | with some friends | from her class | at the gym.

c. She had joined | a softball league | the previous month, | with some friends | from her class | at the gym.

(101) Dwayne rented a car. |

a. He arrived | in Las Vegas | the previous morning, | with a couple | of old friends | from college.

b. He arrived | in Las Vegas | the next morning, | with a couple | of old friends | from college.

c. He had arrived | in Las Vegas | the previous morning, | with a couple | of old friends | from college.

(102) Paige misread a street sign. |

a. She took | a wrong turn | moments earlier, | in her rush | to band rehearsal | at the school.

b. She took | a wrong turn | moments later, | in her rush | to band rehearsal | at the school.

c. She had taken | a wrong turn | moments earlier, | in her rush | to band rehearsal | at the school.

(103) Luisa scrubbed her kitchen counter. |

a. She vacuumed | all the carpets | before that, | in an anxious mood | about the party | at her house.

b. She vacuumed | all the carpets | after that, | in an anxious mood | about the party | at her house.

c. She had vacuumed | all the carpets | before that, | in an anxious mood | about the party | at her house.

(104) Larry failed the midterm exam. |

a. He stayed up | all night worrying | a few days earlier, | at a desk | in the library | by himself.

b. He stayed up | all night worrying | a few days later, | at a desk | in the library | by himself.

c. He had stayed up | all night worrying | a few days earlier, | at a desk | in the library | by himself.

B.4 Experiment 1D: Maze Plausibility

(105) Jeannie raced down the alleyway. She caught the burglar.

a. She found a telltale fingerprint an hour before that, at the scene of the crime with her colleagues.

b. She found a telltale fingerprint an hour after that, at the scene of the crime with her colleagues.

c. She took him to the station an hour before that, at the scene of the crime with her colleagues.

d. She took him to the station an hour after that, at the scene of the crime with her colleagues.

(106) Eddie and Francine started walking. Eddie fell.

a. Francine pushed him from behind a moment earlier, on the path to the lake behind the house.

b. Francine pushed him from behind a moment later, on the path to the lake behind the house.

c. Francine tripped over a loose rock a moment earlier, on the path to the lake behind the house.

d. Francine tripped over a loose rock a moment later, on the path to the lake behind the house.

(107) Maxine went hiking in the woods. She broke her foot.

a. She fell awkwardly on it prior to that, at the hospital down the street from her apartment.

b. She fell awkwardly on it following that, at the hospital down the street from her apartment.

c. She got it checked prior to that, at the hospital down the street from her apartment.

d. She got it checked following that, at the hospital down the street from her apartment.

(108) Felicia and Gerald ran some errands. She criticized him.

a. He bought the wrong bread the day before that, in line at the grocery store near the bank.

b. He bought the wrong bread the day after that, in line at the grocery store near the bank.

c. He nagged her about the chores the day before that, in line at the grocery store near the bank.

d. He nagged her about the chores the day after that, in line at the grocery store near the bank.

(109) Charlene and Howard went to the park. She scolded him.

a. He made an enormous mess right before that, on the path in the middle of the park.

b. He made an enormous mess right after that, on the path in the middle of the park.

c. He walked away quickly right before that, on the path in the middle of the park.

d. He walked away quickly right after that, on the path in the middle of the park.

(110) Phoebe saw Kevin at the office. She yelled at him.

a. He stole her favorite pencil earlier that day, from her desk across the room in history class.

b. He stole her favorite pencil later that day, from her desk across the room in history class.

c. He burst into hysterical tears earlier that day, at his desk across the room in history class.

d. He burst into hysterical tears later that day, at his desk across the room in history class.

(111) Patrick broke Nadya's favorite mug. She forgave him.

a. He apologized profusely to her earlier that day, in the aisle of the train to the city.

b. He apologized profusely to her later that day, in the aisle of the train to the city.

c. He breathed a big sigh of relief earlier that day, in the aisle of the train to the city.

d. He breathed a big sigh of relief later that day, in the aisle of the train to the city.

(112) Daisy arrived at Pete's house. He greeted her.

a. She ran into the room moments before that, with bright eyes and an excited grin on her face.

b. She ran into the room moments after that, with bright eyes and an excited grin on her face.

c. She waved happily to him moments before that, with bright eyes and an excited grin on her face.

d. She waved happily to him moments after that, with bright eyes and an excited grin on her face.

(113) Luke and Ruby went to the soccer game. He hugged her.

a. She scored the winning goal earlier that afternoon, at the park with the dog run near the beach.

b. She scored the winning goal later that afternoon, at the park with the dog run near the beach.

c. She tousled his hair playfully earlier that afternoon, at the park with the dog run near the beach.

d. She tousled his hair playfully later that afternoon, at the park with the dog run near the beach.

(114) Julie and Sean enrolled in student government. She voted for him.

a. He gave an excellent speech the previous week, at the debate in the auditorium at the high school.

b. He gave an excellent speech the next week, at the debate in the auditorium at the high school.

c. He gave an excellent speech the previous week, at the debate in the auditorium at the high school.

d. He gave an excellent speech the next week, at the debate in the auditorium at the high school.

(115) Eliza announced Bob's talk at the conference. She interrupted him.

a. He went over his speech time just before that, in the lecture hall at the university in midtown.

b. He went over his speech time just after that, in the lecture hall at the university in midtown.

c. He ignored her raised hand just before that, in the lecture hall at the university in midtown.

d. He ignored her raised hand just after that, in the lecture hall at the university in midtown.

(116) Zoey got mad at her cat. She chased it.

a. It escaped her apartment immediately before, with an angry 'meow' and its claws and fangs bared.

b. It escaped her apartment immediately after, with an angry meow and its claws and fangs bared.

c. It ran up a huge tree immediately before, with an angry meow and its claws and fangs bared.

d. It ran up a huge tree immediately after, with an angry meow and its claws and fangs bared.

(117) Ned woke up late. He forgot his lunch.

a. He packed soup and a sandwich an hour before, in his rush to the bus stop two blocks away.

b. He packed soup and a sandwich an hour after, in his rush to the bus stop two blocks away.

c. He realized his shirt was on inside out an hour before, in his rush to the bus stop two blocks away.

d. He realized his shirt was on inside out an hour after, in his rush to the bus stop two blocks away.

(118) Paul inherited some money. He bought a new phone.

a. He lost his old phone with the cracked screen the day before, on the trail in the woods behind his house.

b. He lost his old phone with the cracked screen the day after, on the trail in the woods behind his house.

c. He imported the contacts from his old phone the day before, at the table by the window in his kitchen.

d. He imported the contacts from his old phone the day after, at the table by the window in his kitchen.

(119) Rob walked over to Laura's office. He consulted her.

a. She stopped by his office the previous morning, in the suite down the hall from her lab.

b. She stopped by his office the next morning, in the suite down the hall from her lab.

c. She referred him to someone else the previous morning, in the suite down the hall from her lab.

d. She referred him to someone else the next morning, in the suite down the hall from his lab.

(120) Mary vowed she would get revenge on Edwin. She tricked him.

a. She made a very elaborate plan the previous evening, in front of all their friends at the party.

b. She made a very elaborate plan the next evening, in front of all their friends at the party.

c. She discovered he was quite gullible the previous evening, in front of all their friends at the party.

d. She discovered he was quite gullible the next evening, in front of all their friends at the party.

(121) Diane moved out of Lucy's house. She divorced Lucy.

a. Lucy dated another woman earlier that year, in the city by the lake near her hometown.

b. Lucy dated another woman later that year, in the city by the lake near her hometown.

c. Lucy bought a new car earlier that year, in the city by the lake near her hometown.

d. Lucy bought a new car later that year, in the city by the lake near her hometown.

(122) Allison scheduled an interview with Jeff. She hired him.

a. He gave great answers to her questions prior to that, with the team in the offices on the tenth floor.

b. He gave great answers to her questions following that, with the team in the offices on the tenth floor.

c. He got the office next to hers prior to that, with the team in the offices on the tenth floor.

d. He got the office next to hers following that, with the team in the offices on the tenth floor.

(123) Adam met Ellen near the finish line. He consoled her.

a. She lost the race by inches a few minutes before, on the field at the track meet in the next town.

b. She lost the race by inches a few minutes after, on the field at the track meet in the next town.

c. She began crying even harder a few minutes before, on the field at the track meet in the next town.

d. She began crying even harder a few minutes after, on the field at the track meet in the next town.

(124) Nellie and Oliver went on their lunch break. She laughed at him.

a. He slipped on a banana peel a moment before, by the table in the middle of the cafeteria.

b. He slipped on a banana peel a moment after, by the table in the middle of the cafeteria.

c. He stopped talking to her a moment before, by the table in the middle of the cafeteria.

d. He stopped talking to her a moment after, by the table in the middle of the cafeteria.

(125) Emily sat down at her desk. She fixed her laptop.

a. She took a computer repair class earlier that year, in the computer lab at the bookstore on campus.

b. She took a computer repair class later that year, in the computer lab at the bookstore on campus.

c. She bought a new laptop sleeve earlier that year, in the computer lab at the bookstore on campus.

d. She bought a new laptop sleeve later that year, in the computer lab at the bookstore on campus.

(126) Edith went to the garage. She replaced her car tires.

a. She noticed the treads were worn two days before, on the highway in the snow and the wind.

b. She noticed the treads were worn two days after, on the highway in the snow and the wind.

c. She drove safely to work two days before, on the highway in the snow and the wind.

d. She drove safely to work two days after, on the highway in the snow and the wind.

(127) Maya thanked Joseph. He found her dog.

a. It got off its leash earlier that afternoon, on the trail into the woods outside town.

b. It got off its leash later that afternoon, on the trail into the woods outside town.

c. He saw her flyers earlier that afternoon, on the trail into the woods outside town.

d. He saw her flyers later that afternoon, on the trail into the woods outside town.

(128) Kate followed the squirrel. She scared it.

a. It snuck into her garden a moment before, from its perch in the tree by the fence.

b. It snuck into her garden a moment after, from its perch in the tree by the fence.

c. It squeaked loudly at her a moment before, from its perch in the tree by the fence.

d. It squeaked loudly at her a moment after, from its perch in the tree by the fence.

(129) Jimmy took out his cleaning supplies. He cleaned his house.

a. His parents said they would visit him the previous night, with great enthusiasm and no regard for his dismay.

b. His parents said they would visit him the next night, with great enthusiasm and no regard for his dismay.

c. His kids tracked mud through the hall the previous night, with great enthusiasm and no regard for his dismay.

d. His kids tracked mud through the hall the next night, with great enthusiasm and no regard for his dismay.

(130) Lianne and Eric sat down next to each other. She poked him.

a. He fell asleep in biology class a moment earlier, in his seat near the back of the lecture hall.

b. He fell asleep in biology class a moment later, in his seat near the back of the lecture hall.

c. He pinched her on the arm a moment earlier, in his seat near the back of the lecture hall.

d. He pinched her on the arm a moment later, in his seat near the back of the lecture hall.

(131) Gideon took out his kitchen scale. He weighed his dinner.

a. He started a new diet prior to that, with the advice of his doctor in mind.

b. He started a new diet following that, with the advice of his doctor in mind.

c. He added more carrots to his plate prior to that, with the advice of his doctor in mind.

d. He added more carrots to his plate following that, with the advice of his doctor in mind.

(132) Simon and Nicole met at a party. He called her.

a. She gave her number to him prior to that, at a restaurant in the middle of the neighborhood.

b. She gave her number to him following that, at a restaurant in the middle of the neighborhood.

c. She asked him to call her back prior to that, at a restaurant in the middle of the neighborhood.

d. She asked him to call her back following that, at a restaurant in the middle of the neighborhood.

(133) Andrea and Zach went outside. She complimented him.

a. He did a complicated dance move just before that, in the shade on the patio in his yard.

b. He did complicated dance move just after that, in the shade on the patio in his yard.

c. He said he liked her jacket just before that, in the shade on the patio in his yard.

d. He said he liked her jacket just after that, in the shade on the patio in his yard.

(134) Harry pulled over. He abandoned his car.

a. It got stuck in a snowbank an hour earlier, on the side of the highway through the mountains.

b. It got stuck in a snowbank an hour later, on the side of the highway through the mountains.

c. He called a towing service an hour earlier, on the side of the highway through the mountains.

d. He called a towing service an hour later, on the side of the highway through the mountains.

(135) Caleb called Denise to his office. He fired her.

a. She embezzled from his company earlier that week, at her cubicle in the middle of the crowded office.

b. She embezzled from his company later that week, at her cubicle in the middle of the crowded office.

c. She cleaned her messy desk earlier that week, at her cubicle in the middle of the crowded office.

d. She cleaned her messy desk later that week, at her cubicle in the middle of the crowded office.

(136) Omar went home. He changed his clothes.

a. He got splattered with mud earlier that day, near the exhibition of modern art at the museum.

b. He got splattered with mud later that day, near the exhibition of modern art at the museum.

c. He went out on a date with Liz earlier that day, at the exhibition of modern art at the museum.

d. He went out on a date with Liz later that day, at the exhibition of modern art at the museum.

(137) Millie scrunched up her face. She sneezed.

a. She caught a nasty cold right before that, at a busy cafe full of students and professors.

b. She caught a nasty cold right after that, at a busy cafe full of students and professors.

c. She started sniffling miserably right before that, at a busy cafe full of students and professors.

d. She started sniffling miserably right after that, at a busy cafe full of students and professors.

(138) Dwayne went to the car dealership. He rented a car.

a. He wrecked his old car the previous morning, with a couple of old friends from college.

b. He wrecked his old car the next morning, with a couple of old friends from college.

c. He arrived in Las Vegas the previous morning, with a couple of old friends from college.

d. He arrived in Las Vegas the next morning, with a couple of old friends from college.

(139) Paige stopped at a red light. She misread a street sign.

a. She forgot her glasses at home minutes earlier, in her rush to band rehearsal at the school.

b. She forgot her glasses at home minutes later, in her rush to band rehearsal at the school.

c. She took a wrong turn minutes earlier, in her rush to band rehearsal at the school.

d. She took a wrong turn minutes later, in her rush to band rehearsal at the school.

(140) Luisa got a new sponge. She scrubbed her kitchen counter.

a. She spilled some soup before that, in an anxious mood about the party at her house.

b. She spilled some soup after that, in an anxious mood about the party at her house.

c. She vacuumed all the carpets before that, in an anxious mood about the party at her house.

d. She vacuumed all the carpets after that, in an anxious mood about the party at her house.

(141) Larry left the classroom. He failed the midterm exam.

a. He decided not to study for it a few days earlier, at a desk in the library by himself.

b. He decided not to study for it a few days later, at a desk in the library by himself.

c. He stayed up all night a few days earlier, at a desk in the library by himself.

d. He stayed up all night a few days later, at a desk in the library by himself.

(142) Sean and Ashley watched a news program about the drought. He chided her.

a. She left the hose on an hour earlier, among the fruit trees on the side of their yard.

b. She left the hose on an hour later, among the fruit trees on the side of their yard.

c. She sprinkled fertilizer an hour earlier, among the fruit trees on the side of their yard.

d. She sprinkled fertilizer an hour later, among the fruit trees on the side of their yard.

(143) John and Kate started a law firm together. She sued him.

a. He stole some important documents a month before, from the safe in the wall of their office.

b. He stole some important documents a month later, from the safe in the wall of their office.

c. She took out an important document a month before, from the safe in the wall of their office.

d. She took out an important document a month later, from the safe in the wall of their office.

(144) Andrew sent Evelyn a present. He called her.

a. She celebrated her birthday a day earlier, at a party at her home in the suburbs.

b. She celebrated her birthday a day later, at a party at her home in the suburbs.

c. She thanked him for the gift a day earlier, at a party at her home in the suburbs.

d. She thanked him for the gift a day later, at a party at her home in the suburbs.

(145) Alice and Wayne got home from work. She berated him.

a. He backed their car into a mailbox an hour earlier, on the corner of the street behind their office.

b. He backed their car into a mailbox an hour later, on the corner of the street behind their office.

c. He stormed off in a huff an hour earlier, on the corner of the street behind their office.

d. He stormed off in a huff an hour later, on the corner of the street behind their office.

(146) Juan took some groceries out of his car. He slipped.

a. He stepped on an icy patch a moment before, on the sidewalk in front of his house.

b. He stepped on an icy patch a moment after, on the sidewalk in front of his house.

c. He regained his footing a moment before, on the sidewalk in front of his house.

d. He regained his footing a moment after, on the sidewalk in front of his house.

(147) Brandon purchased a haunted mansion in Spain. He called a team of ghost hunters.

a. He saw a ghost the previous morning, near the door from the kitchen to the garage.

b. He saw a ghost the next morning, near the door from the kitchen to the garage.

c. They found evidence of ghosts the previous morning, near the door from the kitchen to the garage.

d. They found evidence of ghosts the next morning, near the door from the kitchen to the garage.

(148) Marisa and Jamal got jobs at a secret government agency. He accused her of treason.

a. She contacted a foreign spy a week earlier, on a secure line from an office in a restricted area.

b. She contacted a foreign spy a week later, on a secure line from an office in a restricted area.

c. She contacted their boss a week earlier, on a secure line from an office in a restricted area.

d. She contacted their boss a week later, on a secure line from an office in a restricted area.

(149) Tori called her parents. She started to cry.

a. Her car was stolen the night before, near a park in a busy part of town.

b. Her car was stolen the night after, near a park in a busy part of town.

c. She met up with a friend the night before, near a park in a busy part of town.

d. She met up with a friend the night after, near a park in a busy part of town.

(150) Clive and Justina went to a friend's house. He pulled a prank on her.

a. She pranked him an hour earlier, in front of all their friends at a bar uptown.

b. She pranked him an hour later, in front of all their friends at a bar uptown.

c. She laughed along with him an hour earlier, in front of all their friends at a bar uptown.

d. She laughed along with him an hour later, in front of all their friends at a bar uptown.

(151) Bruce went to the store. He bought Rogaine.

a. His hair began to thin a few months earlier, from his forehead to the center of his head.

b. His hair began to thin a few months later, from his forehead to the center of his head.

c. His hair began to thicken a few months earlier, from his forehead to the center of his head.

d. His hair began to thicken a few months later, from his forehead to the center of his head.

(152) Olga locked her house up tight. She went to a storm shelter.

a. She heard a hurricane warning an hour earlier, on the radio app on her phone with the cracked screen.

b. She heard a hurricane warning an hour later, on the radio app on her phone with the cracked screen.

c. She listened to a podcast an hour earlier, on the radio app on her phone with the cracked screen.

d. She listened to a podcast an hour later, on the radio app on her phone with the cracked screen.

(153) Kiyana slammed on her brakes. She made a turn.

a. She dropped her wallet an hour before, at a bakery with famous cookies in a trendy neighborhood.

b. She dropped her wallet an hour after, at a bakery with famous cookies in a trendy neighborhood.

c. She met with her boss an hour before, at a bakery with famous cookies in a trendy neighborhood.

d. She met with her boss an hour after, at a bakery with famous cookies in a trendy neighborhood.

(154) Janice opened her closet. She put on her raincoat.

a. She saw dark clouds a few minutes later, near the woods not far from her house in the street.

b. She saw dark clouds a few minutes earlier, near the woods not far from her house in the street.

c. She went for a walk a few minutes later, near the woods not far from her house in the street.

d. She went for a walk a few minutes earlier, near the woods not far from her house in the street.

(155) Leslie got on the highway. She pulled her car over.

a. She heard ambulance sirens seconds later, a short distance down the road from the highway entrance.

b. She heard ambulance sirens seconds earlier, a short distance down the road from the highway entrance.

c. She changed radio stations seconds later, a short distance down the road from the highway entrance.

d. She changed radio stations seconds earlier, a short distance down the road from the highway entrance.

(156) Mahina worked late on a presentation. She woke up early.

a. She had an anxiety dream a few minutes after, on the couch by her desk in the den.

b. She had an anxiety dream a few minutes before, on the couch by her desk in the den.

c. She added the finishing touches a few minutes after, on the couch by her desk in the den.

d. She added the finishing touches a few minutes before, on the couch by her desk in the den.

(157) Alyssa took out her handkerchief. She blew her nose.

a. She inhaled some flour a moment later, in the corner of the pantry with the baking ingredients.

b. She inhaled some flour a moment earlier, in the corner of the pantry with the baking ingredients.

c. She exhaled slowly a moment later, in the corner of the pantry with the baking ingredients.

d. She exhaled slowly a moment earlier, in the corner of the pantry with the baking ingredients.

(158) Nancy was nominated for an economics award. She won it.

a. She did pioneering work the year after, about loans for women in a rural community.

b. She did pioneering work the year before, about loans for women in a rural community.

c. She edited several papers the year after, about loans for women in a rural community.

d. She edited several papers the year before, about loans for women in a rural community.

(159) Barbara ran out of her dorm. She cheered.

a. She got a perfect report card a minute earlier, via the student portal on the university website.

b. She got a perfect report card a minute earlier, via the student portal on the university website.

c. She enrolled in classes a minute earlier, via the student portal on the university website.

d. She enrolled in classes a minute earlier, via the student portal on the university website.

(160) Susan walked out of the kitchen. She spilled her coffee.

a. She bumped into a table a moment earlier, in the corner by the door to the dining room.

b. She bumped into a table a moment later, in the corner by the door to the dining room.

c. She frowned grumpily a moment earlier, in the corner by the door to the dining room.

d. She frowned grumpily a moment later, in the corner by the door to the dining room.

(161) Chantias called her insurance agent. She filed a claim.

a. She got sideswiped in traffic a day earlier, near an intersection two blocks away from her school.

b. She got sideswiped in traffic a day later, near an intersection two blocks away from her school.

c. She got an electronic payment a day earlier, near an intersection two blocks away from her school.

d. She got an electronic payment a day later, near an intersection two blocks away from her school.

(162) Bill hired Michelle as his lawyer. She called him.

a. He missed their appointment a day earlier, at her law office in a new skyscraper in Brooklyn.

b. He missed their appointment a day later, at her law office in a new skyscraper in Brooklyn.

c. She went to a lunch meeting a day earlier, at her law office in a new skyscraper in Brooklyn.

d. She went to a lunch meeting a day later, at her law office in a new skyscraper in Brooklyn.

(163) Saul delivered a speech on paper sales to Shauna. She corrected him.

a. He said something erroneous minutes earlier, at the head of the large table in the conference room.

b. He said something erroneous minutes later, at the head of the large table in the conference room.

c. He took a sip of water minutes earlier, at the head of the large table in the conference room.

d. He took a sip of water minutes later, at the head of the large table in the conference room.

(164) Diyara and Jack bought a fridge for their office. She complained about him.

a. He stole her fruit cup the previous afternoon, in the break room near their workstations by the window.

b. He stole her fruit cup the next afternoon, in the break room near their workstations by the window.

c. He brewed some coffee the previous afternoon, in the break room near their workstations by the window.

d. He brewed some coffee the next afternoon, in the break room near their workstations by the window.

(165) Leilani went to the auto parts store. She bought a tire patch kit.

a. She ran over a nail an hour earlier, in a construction zone on a busy street on her commute.

b. She ran over a nail an hour later, in a construction zone on a busy street on her commute.

c. She asked her cousin for advice an hour earlier, in a construction zone on a busy street on her commute.

d. She asked her cousin for advice an hour later, in a construction zone on a busy street on her commute.

(166) Aidan went out onto the patio. He fired up the grill.

a. He volunteered to barbeque chicken two days earlier, for a potluck at a friend's house down the street.

b. He volunteered to barbeque chicken two days later, for a potluck at a friend's house down the street.

c. He chilled a case of beer two days earlier, for a potluck at a friend's house down the street.

d. He chilled a case of beer two days later, for a potluck at a friend's house down the street.

(167) Candido and Malina finished their sodas. They bought tickets to an indie

movie.

a. They read a good review of it the day before, in the paper with the insert full of ads.

b. They read a good review of it the day after, in the paper with the insert full of ads.

c. They read the weather forecast the day before, in the paper with the insert full of ads.

d. They read the weather forecast the day after, in the paper with the insert full of ads.

(168) Deshaun got out his fishing gear. He went fishing in a hidden cove.

a. He learned about the spot a week earlier, at the tackle shop on the docks in the harbor.

b. He learned about the spot a week later, at the tackle shop on the docks in the harbor.

c. He bought a new reel a week earlier, at the tackle shop on the docks in the harbor.

d. He bought a new reel a week later, at the tackle shop on the docks in the harbor.

Appendix C

Experiment 2 Materials

C.0.1 3A (Narration) Items

(1) Wayne went to the farmers' market. He needed some inspiration to finish his weekly meal-planning. He checked out all the dairy stalls first. He got his favorite cheeses, and some he hadn't tried before. Right after that, he browsed through the produce. First he found some peaches. He chose a few nice ears of corn. He bought a big bunch of parsley on sale.

{A moment/Ten minutes/A day} later, he headed over to the baked goods section.

He decided to invite some friends over for dinner. He was excited to share all the delicious things he'd bought.

Probe: BOUGHT

(2) Luis did a bunch of landscaping work in his garden. First, he installed a border of drought-resistant shrubs. Next, he built some vegetable beds. He moved on to flowers after that. He wanted to attract bees, so he planted poppies and lavender. Then, he placed a birdfeeder near his Japanese maple.

{A moment/Ten minutes/A couple days} later, he filled it to the brim with sunflower seeds.

Last, he arranged his patio furniture facing the tree across the lawn. He relaxed in one of the chairs and waited for some birds to come to the feeder.

Probe: EARNED

(3) Joel was in a foul mood this morning. First, he discovered that his roommates had left a mess in the kitchen. Then he realized he was out of both coffee and cereal. He rushed to a cafe to get breakfast before work. He received a very poorly made caramel latte.

{A moment/Five minutes/A couple days} later, he remembered that he had an early meeting at work.

He had to sprint to catch the train. He made it to the meeting, but arrived quite sweaty and grumpy.

Probe: RECEIVED

(4) Last weekend, Mark realized that he had some unhealthy online habits. On Saturday, he made several impulse buys. He somehow spent two hours that

evening trying to take a perfect selfie. Sunday was no better. He started a commenting war on social media.

{A moment/Five minutes/Several hours} later, he logged onto his usual online poker site.

He lost several hundred dollars before finally logging off. Feeling guilty and irresponsible, he resolved to make better choices online.

Probe: MARRIED

(5) Ryan spent some quality time with his dog Buster this weekend. He first took Buster on a nice long walk in the woods. Then they drove to the pet store. Ryan let Buster pick out a new chew toy. He bought a new brush for Buster as well. On the way home, Ryan let Buster hang his head out the window. Later, he brushed Buster's shaggy coat from head to tail.

{A moment/Ten minutes/A couple days} later, he and Buster settled in to watch a movie and eat popcorn.

Finally, the two of them went upstairs to get ready for bed. Ryan felt recharged and ready to face the new week.

Probe: BRUSHED

(6) Padma is a big science fiction fan. Last Wednesday, she devoted hours to her fan interests. First, she bought a limited edition figurine online. After that, she watched a livestreamed interview with her favorite sci-fi author. Later, she

streamed the first episode of the Mandalorian.

{A moment/Ten minutes/A day} later, she searched for fabric for the cosplay she was working on.

She intended to enter a cosplay contest, so she wanted all the details to be perfect. She just needed a few more materials to complete the outfit.

Probe: SMOOTHED

(7) Arthur did some volunteer work on Saturday. He was part of an environmental justice organization. The group was fighting to get anti-fracking legislation passed. They needed to get several thousand signatures on a petition. Arthur's job was to gather signatures downtown. He asked a passing shopper to sign the petition.

{A moment/Ten minutes/A day} later, a large cluster of college students passed by.

Arthur got them all to sign. By the end of the day, he'd collected an impressive number of signatures.

Probe: HEARD

(8) Isaiah house-sat for his neighbors last week. It was a good way for him to make a little extra money. He made some uncharacteristic mistakes on Wednesday, though. He was working on his laptop in the upstairs study. He knocked his coffee all over the desk and rug.

{A moment/Five minutes/A day} later, he also managed to break a picture frame.

He got everything cleaned up, and bought a replacement frame. He hoped his neighbors wouldn't notice anything amiss.

Probe: LOOKED

(9) Elena is a professional soccer player. She is a big part of the campaign to get women players equal pay to the men. Last Monday, she did an interview with a major news program about the campaign. She praised her teammates and coaches. A couple minutes later, she made some scathing comments about the owners of her team. She criticized their lack of support for their own players.

{A moment/Twenty minutes/Two days} later, she also called out the league executives for undermining the campaign.

In the end, the league fined Elena for a supposed /'conduct violation/'. Elena got the last laugh, though–the move only increased public support for the campaign.

Probe: CRITICIZED

(10) Lorena has a huge sneaker collection. She's always on the lookout for limited edition pairs. On Saturday, she found a very rare pair on an auction website. She quickly placed a bid. The competition was fierce–three more bids topped hers within minutes. She raised her bid by a hundred dollars–four times the minimum bid increase.

{A moment/A half hour/A few hours} later, she got a notification that someone else had placed a bid.

Lorena immediately bid again. The auction was about to end, and she was determined to win.

Probe: FORCED

(11) Kim and DeAndre are best friends. They always go the extra mile to look out for one another. Last weekend was a perfect example of this. Kim offered to help DeAndre move, even though they live in different cities.

{A moment/Ten minutes/A day} later, she also began looking for a great housewarming gift.

In return, DeAndre paid Kim's plane fare to her sister's wedding. He knew she'd been worried about the cost, and didn't want her to miss out.

Probe: OFFERED

(12) Mark is in the midst of planning his wedding. His sister Renee has been helping him. However, on Monday, they got completely fed up with each other. Mark argued heatedly with Renee about the seating arrangements.

{A moment/Twenty minutes/A day} later, he accused her of ordering the wrong color of napkins.

That set Renee off, because it was Mark who had ordered the napkins. She quickly left the house before she said something she'd regret.

Probe: FOLDED

(13) Alicia had the starring role in a superhero movie until last Thursday. That was when she abruptly quit the production. The full story of what went wrong hasn't come out yet. She posted a short statement about staying positive and moving forward.

{A moment/Ten minutes/Several hours} later, she followed that up with a photo of her hugging her costar.

That sparked speculation that she'd quit because of the director, not the other actors. Even with the rumor mill in overdrive, Alicia went silent after that.

Probe: POSTED

(14) Lily did some damage to her bike on Sunday morning. She was on her usual ride through the neighborhood. As she neared home, things took an unexpected turn for the worse. She crashed her bike into a low wall.

{A moment/Ten minutes/A day} later, she stepped in a mud puddle on her own driveway.

When she examined the bike, she found several broken spokes and deep scratches. She took it to the repair shop in town for an overhaul.

Probe: LAUGHED

(15) On Friday, Gail spent a long time cooking dinner. She had a few mishaps that

slowed her down. Straight away, she cut her finger while peeling carrots for a stew. Then she burned some garlic. That set off the smoke alarm. Finally, she got the stew going. She put in meat, vegetables, and broth. Then, she added some corn starch to the stew.

{A moment/Five minutes/A day} later, she also threw in a bundle of herbs.

She left the stew simmering and turned to her next task. She wanted to make an intricate tart with fresh summer fruits.

Probe: ADDED

(16) Neil went on a day trip to a small beach town. He hung out at the beach until lunch time. First, he went for a swim. Then he did some shoreline fishing. After lunch, he went to a shop that sold local products. He examined a display of amusing greeting cards.

{A moment/Five minutes/Five hours} later, he remembered that his sister's birthday was coming up.

He bought a birthday card and a sea glass necklace for his sister. He also got a bag of saltwater taffy to share with his roommates later.

Probe: INCLUDED

(17) On Sunday, Tess did all her outdoor chores. She watered and weeded the vegetable garden. She was careful not to overwater the radishes. Next, she swept the front porch and the back patio. The patio had been covered in au-

tumn leaves. A couple hours later, she did some maintenance on her car. She washed it, then checked the oil. Then she installed a new pair of windshield wipers.

{A moment/Ten minutes/Ten hours} later, she noticed a small leak in one of the front tires.

She wasn't sure if she'd be able to repair it. She called her cousin, a car mechanic, for advice.

Probe: RESIGNED

(18) Sam loves Animal Crossing. Last weekend, he focused on the gardening activities. First he planted some shrubs. A couple minutes later, he bought flower seeds. Then, he pulled all the weeds he could find.

{A moment/Fifteen minutes/Fifteen hours} later, he sold them to a villager for double the normal price.

It took ages to pull all those weeds, but the profits made it well worth the time. Once the sale was complete, Sam began planning how to spend the bounty.

Probe: BEGAN

(19) Kerry explored her new neighborhood on Saturday. She walked several blocks in each direction from her apartment. She found the nearest laundromat. Afer that, she browsed in a mom-and-pop bookstore. Then, she went into a bakery on the corner.

{A moment/Ten minutes/Ten hours} later, she became hungry enough to eat a bear.

The bakery was warm, cozy, and full of delicious-looking treats. Kerry knew she'd become a regular there.

Probe: WENT

(20) Ernest worked hard on Tuesday. He started off with house calls to his most elderly patients. Some of them needed extensive care. Then he spent the afternoon on his feet, doing rounds at the senior care facility. He finally got home well after dark. He collapsed into an overstuffed armchair by the fireplace.

{A moment/Ten minutes/Ten hours} later, he pulled a cozy blanket over his knees.

He knew he should get up and eat dinner, but was too tired. Almost immediately, he was fast asleep.

Probe: WITNESSED

(21) Eileen had an awful day at work yesterday. First, she spilled coffee all over her pants. An hour later, a coworker yelled at her for someone else's mistake. Things didn't improve as the day wore on. She returned from lunch much later than expected.

{A moment/Five minutes/Five hours} later, she remembered that she had a deadline coming up.

She ended up having to stay at the office several hours late. Because of that, she missed out on after-work drinks with her friends.

Probe: RETURNED

(22) Bonnie has a ton of plants in her apartment. Every evening she takes care of them. Last night, she did her rounds as usual. She watered. She trimmed dead leaves and overgrown shoots. Next she focused on the plants that were struggling. First she repotted an orchid that had outgrown its pot. She sprayed her drooping ficus with a special fertilizer.

{A moment/Five minutes/Several hours} later, she moved a sad-looking cactus to a sunnier spot.

For the next week, she anxiously checked on all three plants twice a day. Gradually, they all regained their health.

Probe: HIT

(23) Rachel spent the weekend in her family's mountain cabin. First, she built a nice fire to warm up the building. Then she made a big pot of stew. While the stew cooked, she read a book. Half an hour later, she went to sit by the hearth. She threw some more wood onto the fire.

{A moment/Ten minutes/Many hours} later, she checked to see if the stew was ready.

She had waited too long to stir it–the bottom was pretty burnt. The stew tasted fine, but Rachel was dreading having to clean the pot.

Probe: NOTICED

(24) Jodie spent the day fishing on Saturday. She went to her favorite spot at the lake behind her house. In the morning, she caught a few small perch. She decided to target bigger fish after lunch. At around PM, she tied a new hook onto her line. Then she put fresh bait on the hook.

{A moment/Five minutes/A day} later, she cast towards a deeper part of the lake.

Her strategy worked, and soon she had several large trout in her cooler. She kept one for her dinner, and shared the rest with her neighbors.

Probe: WON

(25) Julie did some major repairs on her niece's teddy bear a few days ago. It was an old bear, and very worn. It needed a complete makeover. First, Julie carefully removed its head and limbs. She pulled out the stuffing, which had taken on a musty smell. She restuffed the bear's arm with cotton batting.

{A moment/Fifteen minutes/A day} later, she replaced its tattered paw pads with new felt.

Julie's progress on the repairs was slow, but she didn't get discouraged. She patiently fixed all the parts, then reassembled the bear.

Probe: RESTUFFED

(26) Last week, Kelly got herself into serious legal trouble. She worked at an upscale jeweler's shop downtown. She stuffed half a dozen diamond bracelets into her bag.

{A moment/Fifteen minutes/Half a day} later, she tucked a bunch of gold earrings into her jacket pockets.

She and her boss hated each other. Kelly knew she'd get fired anyway, so she figured she might as well make it interesting.

Probe: CRUSHED

(27) Olive had a miserable couple of days last week. Her acting career had been steadily declining, but last week was especially bad. First, she found out that the series she was on was going to be cancelled. Her prospects for new jobs weren't looking very good. She learned that she had lost a coveted role to a younger actor.

{A moment/A half hour/A day} later, she got more bad news from her long-time agent, Carolyn.

A few productions were interested in her, but they wanted her to play 'aging mother' roles. Olive was horrified, since she believed she was much too young for such roles.

Probe: LOST

(28) Dan turned 40 last month. Unlike a lot of his friends, he didn't bemoan getting

‘old’ at the time. However, this past week, he discovered that his body is indeed aging. He threw out his back for the first time ever.

{A moment/Five minutes/Many hours} later, he found a cluster of white hairs in his beard.

That felt like adding insult to injury. He lay in bed for the rest of the day, feeling quite sorry for himself.

Probe: STOLE

(29) Nadiya started a new health regimen on Monday. She downloaded an app to track her exercise, diet, and weight. Then she made herself a vitamin-packed smoothie. An hour later, she put on her new workout gear and went to her exercise space. She moved through a yoga routine with relative ease.

{A moment/Fifteen minutes/Two days} later, she attempted to follow a kick-boxing workout video on YouTube.

She had to give up less than halfway through. She felt faint and could barely lift her arms.

Probe: ATTEMPTED

(30) Allison attempted a DIY clock repair project last Thursday. She had found an old wall clock in her grandma’s garage. First, she removed the mechanism from the wooden case. She wound it up carefully. To her surprise, it began ticking right away. It wasn’t working smoothly, though. She kept tinkering

with it. She loosened a screw just a tiny bit.

{A moment/Ten minutes/Several hours} later, she brushed some dust off of a tiny spring.

Patiently, Allison kept making little adjustments and cleaning the clock parts. When she was satisfied that the clock was keeping good time, she put it back in its case.

Probe: LOOSENED

C.0.2 Items appearing in 3B (Result) and 3C (Explanation)

(31) Wayne went to the farmers' market. He needed some inspiration to finish his weekly meal-planning. He checked out all the dairy stalls first. He got his favorite cheeses, and some he hadn't tried before. Right after that, he browsed through the produce. First he found some peaches. He chose a few nice ears of corn. He bought a big bunch of parsley on sale.

{A moment/Ten minutes/A day} {earlier/later}, he had/$\emptyset$ decided to make ham with parsley sauce for dinner.

He invited some friends over for an impromptu get-together. He asked them to bring some side dishes and drinks.

Probe: BOUGHT; Comp Q (easy): Did Wayne buy cheese?

(32) Luis did a bunch of landscaping work in his garden. First, he installed a border

of drought-resistant shrubs. Next, he built some vegetable beds. He moved on to flowers after that. He wanted to attract bees, so he planted poppies and lavender. Then, he placed a birdfeeder near his Japanese maple.

{A moment/Ten minutes/A couple days} {earlier/later}, he had/∅ noticed an adorable bird land atop the maple.

Last, he arranged his patio furniture facing the tree across the lawn. He relaxed in one of the chairs and watched as more birds came by the feeder.

Probe: EARNED; Comp Q (easy): Does Luis have a garden?

(33) Heather had to deal with a couple of health problems last week. On Monday, she had an awful toothache. She went to the dentist to get it taken care of. A couple days later, another problem arose. She broke out in a rash all over her arms.

{A moment/Five minutes/A day} {earlier/later}, she had/∅ put on lotion before going to bed.

She rinsed her arms thoroughly, but the rash remained. She began searching for a dermatologist online.

Probe: BROKE; Comp Q (easy): Was Heather completely healthy last week?

(34) Emi is a serious amateur tennis player. She practices every morning before work. She's always fine-tuning her technique. Her coach gave her a few tips about serving last week. First, they worked out the new motions together. The

next day, Emi went to the court on her own. She felt her right wrist give a slight twinge.

{A moment/Ten minutes/A day} {earlier/later}, she had/$\emptyset$ changed how she gripped the raquet.

She served another bucket of balls, trying to follow her coach's advice. Eventually, she figured out the new motion, and her wrist began to feel better.

Probe: SENT; Comp Q (easy): Does Emi play squash?

(35) Joel was in a foul mood this morning. First, he discovered that his roommates had left a mess in the kitchen. Then he realized he was out of both coffee and cereal. He rushed to a cafe to get breakfast before work. He received a very poorly made caramel latte.

{A moment/Five minutes/A couple days} {earlier/later}, he had/$\emptyset$ shouted at the barista in front of everyone.

The day was almost over by the time he finally cooled off. He resolved to apologize to the barista the next morning.

Probe: RECEIVED; Comp Q (hard): Did Joel's roommates do something inconsiderate?

(36) Last weekend, Mark realized that he had some unhealthy online habits. On Saturday, he made several impulse buys. He somehow spent two hours that evening trying to take a perfect selfie. Sunday was no better. He started a

commenting war on social media.

{A moment/Five minutes/Several hours} {earlier/later}, he had/∅ gotten angry about a rude, anonymous tweet.

He didn't log off until the middle of the night. He called his sister to vent, but her groggy responses finally made him see that he had a problem.

Probe: MARRIED; Comp Q (hard): Did Mark start a commenting war after he made some impulse purchases?

(37) Ryan spent some quality time with his dog Buster this weekend. He first took Buster on a nice long walk in the woods. Then they drove to the pet store. Ryan let Buster pick out a new chew toy. He bought a new brush for Buster as well. On the way home, Ryan let Buster hang his head out the window. Later, he brushed Buster's shaggy coat from head to tail.

{A moment/Ten minutes/A couple days} {earlier/later}, Buster had begun shedding clumps of fur all over the house.

After Buster was all groomed, Ryan gave him some treats. He swept up all the fur, then sat with Buster and watched a movie.

Probe: BRUSHED; Comp Q (hard): Did Ryan and Buster go for walk after going to the pet store?

(38) Padma is a big science fiction fan. Last Wednesday, she devoted hours to her fan interests. First, she bought a limited edition figurine online. After that,

she watched a livestreamed interview with her favorite sci-fi author. Later, she streamed the first episode of the Mandalorian.

{A moment/Ten minutes/A day} {earlier/later}, she had/∅ gotten into all things involving Baby Yoda.

She ended up binge-watching the entire first season. Then she browsed online for a Baby Yoda figurine to add to her collection.

Probe: SMOOTHED; Comp Q (hard): Does Padma think figurines are a waste of money?

(39) Lizzie had a pet-related fiasco this past weekend. She was looking after her sister's cat, Nigel. Unfortunately, she and Nigel hated each other. Saturday passed in relative peace. The real trouble came on Sunday. Lizzie cornered Nigel, who yowled and hissed angrily.

{A moment/Five minutes/A few hours} {earlier/later}, Nigel had/∅ escaped from the house into the front yard.

Eventually, Lizzie wrangled Nigel back inside. Nigel, however, left plenty of scratches on Lizzie's arms.

Probe: CORNERED; Comp Q (easy): Did Nigel get out of the house?

(40) Arthur did some volunteer work on Saturday. He was part of an environmental justice organization. The group was fighting to get anti-fracking legislation passed. They needed to get several thousand signatures on a petition. Arthur's

job was to gather signatures downtown. He asked a passing shopper to sign the petition.

{A moment/Ten minutes/A day} {earlier/later}, the shopper had/∅ given him a friendly smile as she passed by.

By the end of the day, Arthur had done pretty well. He had collected thirty signatures, even though there hadn't been very many people out shopping.

Probe: HEARD; Comp Q (easy): Did Arthur go downtown?

(41) Johnny and his son Kevin made a bit of a scene at the family reunion last weekend. Everyone had gathered at Johnny's house. Johnny and the other adults started off with snacks and drinks in the backyard. Meanwhile, Kevin and all his cousins played all over the house. The drama started just after lunchtime. Johnny scolded Kevin in front of the whole family.

{A moment/Ten minutes/Several hours} {earlier/later}, Kevin had/∅ broken a picture frame with his slingshot.

Their relatives weren't sure how to react. For the most part, they tried to pretend nothing was wrong.

Probe: SCOLDED; Comp Q (easy): Did Kevin break a vase?

(42) Isaiah house-sat for his neighbors last week. It was a good way for him to make a little extra money. Things took an odd turn on Wednesday, though. Isaiah was working on his laptop in the upstairs study. He turned on the air

conditioner in the corner of the room.

{A moment/Five minutes/A day} {earlier/later}, a weird smell had/∅ begun to fill the room.

He checked the air vents, the drawers, and under the chair cushions for anything out of place. But, in the end, he couldn't figure out what the smell was, or where it was coming from.

Probe: LOOKED; Comp Q (easy): Did Isaiah find the source of the weird smell?

(43) Elena is a professional soccer player. She is a big part of the campaign to get women players equal pay to the men. Last Monday, she did an interview with a major news program about the campaign. She praised her teammates and coaches. A couple minutes later, she made some scathing comments about the executives in charge of her team. She criticized the executives' lack of support for their own players.

{A moment/Twenty minutes/Two days} {earlier/later}, the team's front office had/∅ fined her $5000 for a supposed 'conduct violation'.

They wouldn't say what exactly Elena had done wrong. That made a lot of people suspect that they were trying to thwart the equal pay campaign.

Probe: CRITICIZED; Comp Q (easy): Did Elena do an interview on the news?

(44) Lorena has a huge sneaker collection. She's always on the lookout for limited

edition pairs. On Saturday, she found a very rare pair on an auction website. She quickly placed a bid. The competition was fierce–three more bids topped hers within minutes. She raised her bid by a hundred dollars–four times the minimum bid increase.

{A moment/A half hour/A few hours} {earlier/later}, another bidder had/∅ increased the bid by fifty dollars.

Lorena refreshed the auction page every few minutes. She wanted to know right away if anyone topped her bid.

Probe: FORCED; Comp Q (easy): Was Lorena bidding on sneakers?

(45) Kim and Deshaun are best friends. They always go the extra mile to look out for one another. Last weekend was a perfect example of this. Kim offered to help Deshaun move, even though they live in different cities.

{A moment/Ten minutes/A day} {earlier/later}, he had/∅ loaned her the money to fly to her sister's wedding.

Neither of them has biological siblings, so they're especially glad to have each other. They each think of the other as 'found family.'

Probe: OFFERED; Comp Q (easy): Do Kim and Deshaun live in the same city?

(46) Emmy learned a lot about dog training last weekend. More accurately, she learned from her mistakes. Her new dog, Walter, wasn't very well-behaved.

She shut Walter in the bedroom by himself.

{A moment/A half hour/A couple days} {earlier/later}, Walter had/$\emptyset$ chewed up a bunch of shoes.

In the bedroom, Walter gnawed on the furniture until Emmy let him out. She consulted a ton of dog-training websites to find a better way to handle him.

Probe: LOST; Comp Q (easy): Was Walter in the kitchen by himself?

(47) This past weekend, Sean went to the gym a few miles from his house. He had just signed up for a membership as part of his new year's resolution. He realized that his cardiovascular fitness was pretty bad.

{A moment/Five minutes/A day} {earlier/later}, he had/$\emptyset$ failed to jog just a quarter mile on the treadmill without stopping.

He was dismayed, but not discouraged. After that, he struggled through some weight-training, but still left feeling optimistic about his fitness journey.

Probe: REALIZED; Comp Q (hard): Does Sean want to get in better shape?

(48) Mark is in the midst of planning his wedding. His sister Renee has been helping him. However, on Monday, they got completely fed up with each other. Mark argued heatedly with Renee about the seating arrangements.

{A moment/Twenty minutes/A day} {earlier/later}, he had/$\emptyset$ assigned her to the same table as their most annoying cousins.

Renee went out for a walk before the fight could get out of hand. She didn't think she should have to sit with those cousins, after all the help she'd given Mark.

Probe: FOLDED; Comp Q (hard): Is Mark going to get married?

(49) Alicia had the starring role in a superhero movie until last Thursday. That was when she abruptly quit the production. The full story of what went wrong hasn't come out yet. She posted a short statement about staying positive and moving forward.

{A moment/Ten minutes/Several hours} {earlier/later}, the online rumor mill had/∅ gone into overdrive about a possible scandal.

Her words did little to slow down the public speculation. Hollywood beat reporters scrambled to get the real scoop.

Probe: POSTED; Comp Q (hard): Did everyone expect Alicia to quit the movie?

(50) Lily did some damage to her bike on Sunday morning. She was on her usual ride through the neighborhood. As she neared home, things took an unexpected turn for the worse. She crashed her bike into a low wall.

{A moment/Ten minutes/A day} {earlier/later}, a spoke in the front wheel had/∅ snapped.

Luckily, she wasn't badly injured, but her bike was pretty banged up. She had to take it to the repair shop in town.

Probe: LAUGHED; Comp Q (hard): Did Lily crash her bike before a spoke in the front wheel snapped?

C.0.3 Items exclusive to 3B

(51) Eileen had an awful day at work yesterday. First, she spilled coffee all over her pants. An hour later, a coworker yelled at her for someone else's mistake. Things didn't improve as the day wore on. She returned from lunch much later than expected.

{A moment/Five minutes/Five hours} later, she received a scolding email from her boss.

She ended up having to stay at the office three hours late. Because of that, she missed out on after-work drinks with her friends.

Probe: RETURNED; Comp Q (easy): Did Eileen's boss email her?

(52) Bonnie has a ton of plants in her apartment. Every evening she takes care of them. Last night, she did her rounds as usual. She watered. She trimmed dead leaves and overgrown shoots. Then she focused on the plants that were struggling. First she repotted an orchid that had outgrown its pot. Then, she saw that her ficus looked droopy and pathetic.

{A moment/Five minutes/Several hours} later, she sprinkled some new plant food into its soil.

For the next week, she anxiously checked on it twice a day. It was her favorite plant, so she was desperate to keep it alive.

Probe: HIT; Comp Q (easy): Did Bonnie give her ficus new plant food?

(53) Rachel spent the weekend in her family's mountain cabin. First, she built a nice fire to warm up the building. Then she made a big pot of stew. While the stew cooked, she read a book. Half an hour later, she went to sit by the hearth. She noticed that the fire was almost out.

{A moment/Ten minutes/Many hours} later, she threw a large amount of wood onto it.

She watched the flames carefully until she was sure they wouldn't go out. Then she returned to her book.

Probe: NOTICED; Comp Q (easy): Was Rachel at a beach house?

(54) Jodie spent the day fishing on Saturday. She went to her favorite spot at the lake behind her house. In the morning, she caught a few small perch. She decided to target bigger fish after lunch. At around PM, she tied a new hook onto her line. Then she put fresh bait on the hook.

{A moment/Five minutes/A day} later, she caught a large trout in the deepest part of the lake.

When it got dark, she trekked back to the house, her cooler full of fish. For dinner that night, she pan-fried a trout with potatoes and asparagus.

Probe: WON; Comp Q (easy): Did Jodie cook a salmon for dinner?

(55) Julie did some major repairs on her niece's teddy bear a few days ago. It was an old bear, and very worn. It needed a complete makeover. First, Julie carefully removed its head and limbs. She pulled out the stuffing, which had taken on a musty smell. She restuffed the bear's arm with cotton batting.

{A moment/Fifteen minutes/A day} later, a seam burst at the end of its paw.

Julie's progress on the repairs was slow, but she didn't get discouraged. She patiently fixed all the parts, then reassembled the bear.

Probe: RESTUFFED; Comp Q (easy): Did Julie take out the bear's old stuffing?

(56) Last week, Kelly got herself into serious legal trouble. She worked at an upscale jeweler's shop downtown. She stuffed half a dozen diamond bracelets into her bag.

{A moment/Fifteen minutes/Half a day} later, she was fired and quickly escorted off the premises.

The jeweler filed a lawsuit against her. He said she had damaged the bracelets.

Probe: CRUSHED; Comp Q (easy): Did Kelly get fired?

(57) Olive had a miserable couple of days last week. Her acting career had been steadily declining, but last week was especially bad. First, she found out that

the series she was on was going to be cancelled. Her prospects for new jobs weren't looking very good. She learned that she had lost a coveted role to a younger actor.

{A moment/A half hour/A day} later, she fired her longtime agent, a respected Hollywood power player.

She began calling up old producer friends herself to see if they had anything for her. Unfortunately, she was too proud to accept the 'aging mother' roles they offered up.

Probe: LOST; Comp Q (easy): Does Olive want to play 'motherly' types?

(58) Dan turned 40 last month. Unlike a lot of his friends, he didn't bemoan getting 'old' at the time. However, this past week, he discovered that his body is indeed aging. He threw out his back for the first time ever.

{A moment/Five minutes/Many hours} later, he twisted awkwardly while trying to put on a sweater.

That made his back hurt even more. He had to lie in bed trying not to move for the rest of the day.

Probe: STOLE; Comp Q (easy): Did Dan just turn 25?

(59) Nadiya started a new health regimen on Monday. She downloaded an app to track her exercise, diet, and weight. Then she made herself a vitamin-packed smoothie. An hour later, she put on her new workout gear and went to her exer-

cise space. She attempted to follow a kickboxing workout video on YouTube.

{A moment/Fifteen minutes/Two days} later, she bemoaned her lack of physical fitness.

Her arms felt like they each weighed a ton. However, she knew it would pay off if she kept at it.

Probe: ATTEMPTED; Comp Q (hard): Did Nadiya have a smoothie before working out?

(60) Allison attempted a DIY clock repair project last Thursday. She had found an old wall clock in her grandma's garage. First, she removed the mechanism from the wooden case. She wound it up carefully. To her surprise, it began ticking right away. It wasn't working smoothly, though. She kept tinkering with it. She loosened a screw just a tiny bit.

{A moment/Ten minutes/Several hours} later, the mechanism began to emit a high-pitched sqeak.

Patiently, Allison kept making little adjustments until the squeaking subsided. When she was satisfied that the clock was keeping good time, she put it back in its case.

Probe: WITHHELD; Comp Q (hard): Does Allison's clock run on batteries?

C.0.4 Items exclusive to 3C

(61) On Friday, Gail spent a long time cooking dinner. She had a few mishaps that slowed her down. Straight away, she cut her finger while peeling carrots for a stew. Then she burned some garlic. That set off the smoke alarm. Finally, she got the stew going. She put in meat, vegetables, and broth. Then, she added some corn starch to the stew.

{A moment/Five minutes/A day} earlier, she had realized that it was the wrong consistency.

She fixed the stew, but forgot the steamed broccoli. It was a horrible green mush by the time she remembered it.

Probe: ADDED; Comp Q (easy): Did Gail set off a smoke alarm?

(62) Neil went on a day trip to a small beach town. He hung out at the beach until lunch time. First, he went for a swim. Then he did some shoreline fishing. After lunch, he went to a shop that sold local products. He examined a display of amusing greeting cards.

{A moment/Five minutes/A few hours} earlier, he had remembered that his sister's birthday was coming up.

He bought a birthday card and a sea glass necklace for his sister. He also got a bag of saltwater taffy to share with his roommates later.

Probe: INCLUDED; Comp Q (easy): Did Neil go fishing?

(63) Thom had an eventful day at work. There was a fire drill right when he arrived. Then, he gave a presentation to the CEO herself. His team treated him to a fancy lunch after that. Things didn't slow down in the afternoon. He called a meeting with the entire team.

{A moment/Twenty minutes/A few hours} earlier, he had discovered a design flaw in their prototype.

They discussed how to fix the problem. Once they decided on a solution, they had to rework the project schedule.

Probe: CALLED; Comp Q (easy): Was Thom's prototype perfect?

(64) On Sunday, Tess did all her outdoor chores. She watered and weeded the vegetable garden. She was careful not to overwater the radishes. Next, she swept the front porch and the back patio. The patio had been covered in autumn leaves. A couple hours later, she did some maintenance on her car. She washed it, then checked the oil. She then measured the tire pressure more attentively than usual.

{A moment/Ten minutes/Ten hours} earlier, she had noticed a small leak in one of the front tires.

She wasn't sure if she'd be able to repair it. She called her cousin, a car mechanic, for advice.

Probe: RESIGNED; Comp Q (easy): Is Tess's cousin an electrician?

(65) Sam loves Animal Crossing. Last weekend, he focused on the gardening activities. First he planted some shrubs. A couple minutes later, he bought flower seeds. Then, he began pulling all the weeds he could find.

{A moment/Fifteen minutes/A day} earlier, a villager had offered to buy weeds from him for double the normal price.

It took ages to pull all those weeds, but the profits made it well worth the time. Once the sale was complete, Sam began planning how to spend the bounty.

Probe: BEGAN; Comp Q (easy): Did Sam sell his weeds for a good price?

(66) Last weekend, Justin had some life-changing experiences. He had been invited to do a guest performance at a big awards show. A good friend of his was one of the main acts. They had a perfect dress rehearsal first. Then, Justin was introduced to a dizzying array of famous musicians and producers. A couple hours later, he also met a group of record label executives. He received an offer for a big record deal.

{A moment/Ten minutes/A day} earlier, he had delivered a show-stealing performance that electrified the audience.

News of the deal hit social media immediately. However, Justin's friend advised him to save the decision for later, and whisked him off to an after-party.

Probe: FOLLOWED; Comp Q (easy): Did Justin meet a bunch of famous people?

(67) Nellie had to do some emergency plumbing work this weekend. First, she snaked her shower drain. It had become clogged with hair. Then she had to deal with the sink. She shut off the water to the whole house.

{A moment/Five minutes/A couple days} earlier, she had noticed water leaking from the base of the sink.

Once the water was off, she mopped up the leak. Then she tried to figure out what was wrong with the pipes.

Probe: REMOVED; Comp Q (easy): Did Nellie repair her dishwasher this weekend?

(68) On Thursday, Lottie had the house to herself overnight. Her housemates were out of town. She was excited to have some peace and quiet. However, she was also a little nervous to be all alone at night. First, she made sure the doors were locked. Then, she closed and latched all the windows. She flipped on the lights on the back patio.

{A moment/Five minutes/Several hours} earlier, she had heard a weird scratching noise in the yard.

The lights revealed a young raccoon. It was trying to open the trash bin.

Probe: CLIMBED; Comp Q (easy): Was there a possum in Lottie's yard?

(69) Kerry explored her new neighborhood on Saturday. She walked several blocks in each direction from her apartment. She found the nearest laundromat. Afer

that, she browsed in a mom-and-pop bookstore. Then, she went into a bakery on the corner.

{A moment/Ten minutes/A day} earlier, she had become extremely hungry from all the walking around.

The bakery was warm, cozy, and full of delicious-looking treats. Kerry knew she'd become a regular there.

Probe: WENT; Comp Q (hard): Does Kerry like baked goods?

(70) Ernest worked hard on Tuesday. He started off with house calls to his most elderly patients. Some of them needed extensive care. Then he spent the afternoon on his feet, doing rounds at the senior care facility. He finally got home well after dark. He collapsed into an overstuffed armchair by the fire.

{A moment/Ten minutes/Half a day} earlier, he had felt a wave of pain travel up his spine.

He gingerly adjusted his posture and waited for the pain to pass. He was used to getting back pain after working long hours.

Probe: WITNESSED; Comp Q (hard): Did Ernest spend most of his work day sitting down?

C.0.5 Fillers

(71) Katie spent the afternoon tidying up her apartment. First, she dusted all the furniture. After that, she vacuumed everywhere. She disinfected all the surfaces in the bathroom. She even scrubbed the towel bars and the doorknob. Then, she got to work on her messy kitchen. She washed all the dishes in the sink. She put on latex gloves while bleaching the sink.

A moment earlier, she had felt an itchy, burning sensation all over her hands.

The gloves helped a lot, but her hands were still a bit itchy later. She washed them thoroughly with mild handsoap.

Probe: SCRUBBED; Comp Q (easy): Does Katie live in an apartment?

(72) Phoebe had a few ups and downs on Thursday. Things went pretty well at school. She got an A on her math test. At lunch, the cafeteria had her favorite cookies. A few hours later, though, her day took a turn. She sang a simple passage quite poorly during her singing lesson.

A moment later, she got really frustrated with how she sounded.

She even thought about quitting singing altogether. That night, she could barely focus on her homework.

Probe: TUBE; Comp Q (easy): Does Phoebe take singing lessons?

(73) Diana worked from home on Friday. She started the day with a full breakfast.

She never got to do that normally. She had a bit of a setback after that, though. She got quite upset with herself.

Ten minutes earlier, she had dropped a mug and it shattered everywhere.

It was hard for her to recalibrate once she started feeling so frustrated. Eventually, though, she recovered some of the positivity she had started the day with.

Probe: BREAKFAST; Comp Q (easy): Did Diana drop a plate?

(74) Celia and Peter got invited to a dinner party. The host asked them each to bring something to share. They decided to bring things that paired well together. They also planned to place online orders that they could pick up on the way to the party. Celia chose a fine wine as her contribution.

Ten minutes later, Peter ordered a fancy charcuterie board from an upscale grocer.

When they went out to pick up their orders, Celia also bought a box of Belgian chocolates. They were a gift especially for the host.

Probe: RULES; Comp Q (easy): Did Celia buy macarons?

(75) On Thursday, Marcia and Fred got invited to their niece's birthday party. They decided to choose presents that encouraged creativity. That evening, they went online to order the gifts, since they didn't want to risk shopping in-person. Marcia directed Fred to their local art supply store's website. She purchased a sturdy, wooden, child-sized easel.

A few hours earlier, Fred had told her that he was going to buy a set of paints. Everything arrived just in time for the party. Marcia and Fred wrapped it all up, then headed to their niece's house to celebrate with the rest of their pod.

Probe: LOCAL; Comp Q (hard): Do Marcia and Fred think that painting encourages creativity?

(76) Last weekend, Travis dogsat for a friend of his. He had been dreading it all week. The dog, Rufus, was pretty high-maintenance. On Saturday, Travis couldn't get Rufus to eat anything. Later, he had to let Rufus out to pee twice in the middle of the night. Sunday was no better. He gave an order to Rufus in an angry voice.

Two days later, Rufus ran away from him at the dog park.

He got some dirty looks for speaking so harshly to a cute dog like Rufus. He didn't care–none of those people knew what a handful Rufus really was.

Probe: GAME; Comp Q (hard): Did Travis go to the dog park on Sunday?

(77) Hayley recorded a song for her YouTube channel on Friday. First, she laid down the guitar tracks. Then, she did the vocals. It took her a couple hours to get all the harmonies just right. After that, she sat down to do the editing and mixing. She had a few technical difficulties, however. She updated her operating system and rebooted the computer.

A moment earlier, her music editing program had stopped working all of a sudden.

It took her nearly an hour to get back to her project. She had to download and install multiple updates for the editing program to get it up and running again.

Probe: EDITING; Comp Q (hard): Did Hayley get all the harmonies right on the first try?

(78) Pedro spent all day on Wednesday cooped up in his home office. He had to finish writing a big article about a college cheating scandal. In the morning, he wrote several paragraphs very quickly. Feeling accomplished, he rewarded himself with a snack break. When he returned, the words didn't come as easily. He started to feel a bit restless. He opened the windows on either side of his desk.

A moment later, the room became uncomfortably hot and muggy.

Fidgety and increasingly sweaty, he kept trying to write. After an hour of making no progress, Pedro finally gave up and went for a walk to clear his head.

Probe: QUOTATIONS; Comp Q (hard): Does Pedro's office have only one window?

(79) Shannon and her brother Eddie got in trouble for fighting yesterday. They were visiting their grandparents. Their parents had told them to behave themselves. However, Shannon and Eddie could never resist some roughhousing. Neither

of them would let a provocation go unanswered, either. Shannon poked Eddie mercilessly in the ribs.

Five minutes earlier, he had pushed her into the hedges beside the path.

They continued battling until their grandmother caught them pulling each other's hair. She sent them both to their rooms for the entire afternoon.

Probe: BEHAVE; Comp Q (easy): Are Shannon and Eddie siblings?

(80) Helen had an awful couple of days last week. She was laid up with the flu. The first day, she managed to sleep relatively well. She also drank a ton of water and even ate some soup. A day later, she felt way worse. She couldn't bring herself to eat or drink anything. She couldn't sleep either, just toss and turn miserably. In the afternoon, she started to feel even more feverish and uncomfortable.

Ten minutes later, she dragged yet another heavy fleece blanket over herself.

Her throat ached, but she managed to swallow some medicine. It took ages to start working.

Probe: SMART; Comp Q (easy): Did Helen have a sore throat?

(81) Ben and his sister Charlotte got grounded last week. They're very keen on pranking each other. Last week, things got out of hand. Ben tattled on Charlotte to their parents at the breakfast table.

A few hours earlier, she had squeezed a whole bottle of sunscreen into his new backpack.

She hadn't realized that his homework–now ruined–was already in the front pocket. Their parents finally had to put a stop to the pranks.

Probe: KEEN; Comp Q (easy): Did Charlotte put confetti in Ben's backpack?

(82) Felix is the starting center for his college basketball team. He loves his teammates, but has never gotten along well with the coach. The two of them disagree, and even argue, pretty regularly. Things really came to a head last Wednesday. Felix yelled obscenities at the coach in the locker room.

A day later, the coach benched him for the next three games.

Felix's teammates tried to talk the coach out of it, but didn't succeed. Without Felix, the team lost all three games.

Probe: TOLERATE; Comp Q (easy): Do Felix and his coach get along well?

Appendix D

Experiment 3 Materials

D.1 Experiment 3A: Animacy x RFC

(1) a. Jessie saw an elderly man in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. {The man/He} fell over. Jessie circled back to investigate what had happened.

b. Jessie saw a rickety mailbox in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. {The mailbox/It} fell over. Jessie circled back to investigate what had happened.

Comprehension question (easy): Was Jessie on a morning walk?

(2) a. Albert saw a nurse at the bus stop on his way to work. He sat down to wait for the bus. {The nurse/She} dropped her phone. Albert picked it up for

her.

b. Albert saw a newspaper at the bus stop on his way to work. He sat down to wait for the bus. {The newspaper fell off the bench. Albert picked it up to read.

Comprehension question (easy): Did Albert wait for the bus?

(3) a. Beth saw a firefighter in the grocery store yesterday afternoon. She made her way down the dairy aisle. {The firefighter/He} put a gallon of milk in his cart. Beth stopped to peruse the yogurt selection.

b. Beth saw an unattended cart in the grocery store yesterday afternoon. She made her way down the dairy aisle. {The cart/It} rolled into a nearby shelf. Beth stopped to straighten it out again.

Comprehension question (easy): Did Beth go down the dairy aisle?

(4) a. Curtis saw a tall woman at the beach last weekend. He laid out his towel on the sand. {The woman/She} frowned at him. Curtis was too close to the beach volleyball court.

b. Curtis saw a tall sand castle at the beach last weekend. He laid out his towel on the sand. {The sand castle/It} started to collapse. Curtis had accidentally hit it with his towel.

Comprehension question (easy): Did Curtis lay out a towel?

(5) a. Diana saw a butcher at the farmer's market on Wednesday. She stopped at a couple of farm stalls. {The butcher/He} waved her over. Diana went to say hello.

b. Diana saw an enormous squash at the farmer's market on Wednesday. She stopped at a couple of farm stalls. {The squash/It} tipped over. Diana quickly pushed it upright.

Comprehension question (easy): Was Diana at the farmer's market?

(6) a. Edwin saw a cheerleader in the hall in between classes. He shut his locker and headed to class. {The cheerleader/She} smiled at him. Edwin smiled awkwardly back.

b. Edwin saw a flyer in the hall in between classes. He shut his locker and headed to class. {The flyer/It} distracted him. Edwin stopped to read it.

Comprehension question (easy): Did Edwin head to class?

(7) a. Fatima saw a surgeon on her visit to the hospital. She got into the elevator. {The surgeon/He} called after her. He asked her to hold the door.

b. Fatima saw an IV stand on her visit to the hospital. She got into the elevator. {The IV stand/It} rolled after her. She had snagged it with her backpack strap.

Comprehension question (easy): Did Fatima visit the hospital?

(8) a. Gregory saw a new librarian during his most recent weekly trip to the library. He turned in the books he'd finished. {The librarian/She} placed an armload of books on a cart. Gregory nodded at her as he passed.

b. Gregory saw an intriguing book cover during his most recent weekly trip to the library. He turned in the books he'd finished. {The book cover/It} featured a strange-looking script. Gregory opened the book to read the inside jacket.

Comprehension question (easy): Did Gregory turn in some library books?

(9) a. Holly saw a young pilot at the airport the other day. She arrived at her gate and sat down. {The pilot/He} hurried past. Holly figured he was running late.

b. Holly saw a private jet at the airport the other day. She arrived at her gate and sat down. {The jet/It} taxied past the window. Holly wondered who it belonged to.

Comprehension question (easy): Was Holly at the airport?

(10) a. Isaac saw a very busy waitress at his favorite diner. He ordered his usual lunch at the counter. {The waitress/She} hustled by with a huge tray. The diner was packed, but understaffed.

b. Isaac saw a lopsided stool at his favorite diner. He ordered his usual lunch at the counter. {The stool/It} wobbled ominously next to him. Isaac managed to fix it by tightening the seat.

Comprehension question (easy): Did Isaac order lunch?

(11) a. Jade saw a baseball player on the train a couple nights ago. She found a spot to stand near the doors. {The baseball player/He} chuckled at something on his phone. Jade only recognized him because her son was a baseball fanatic.

b. Jade saw a loose baseball on the train a couple nights ago. She found a spot to stand near the doors. {The baseball/It} began rolling around when the train moved. Jade tried to stop it with her foot, but missed.

Comprehension question (easy): Was Jade in a taxi?

(12) a. Kelvin saw a florist while he was walking home on Monday. He adjusted the strap of his messenger bag. {The florist/She} ran up to him. His gloves had fallen out of his bag.

b. Kelvin saw a lovely bouquet while he was walking home on Monday. He adjusted the strap of his messenger bag. {The bouquet/It} featured irises, his favorite flower. It was expensive, but he just had to buy it.

Comprehension question (easy): Was Kelvin riding a motorcycle?

(13) a. Louise saw a groundskeeper at the golf course last weekend. She hefted her golf bag over her shoulder. {The groundskeeper/He} tossed some

fertilizer over a brown patch of grass. Louise greeted him on her way to the next hole.

b. Louise saw a new sand trap at the golf course last weekend. She hefted her golf bag over her shoulder. {The sand trap/It} had replaced a brown patch of grass. Louise inspected it on her way to the next hole.

Comprehension question (easy): Was Louise at the basketball courts last weekend?

(14) a. Morris saw a stylish woman during his morning run through the city. He paused at a stoplight, jogging in place. {The woman/She} hailed a cab. Morris wished his own fashion sense were as good as hers.

b. Morris saw a flashy motorcycle during his morning run through the city. He paused at a stoplight, jogging in place. {The motorcycle/It} roared through the intersection. Morris wished his own motorcycle looked as good as that one.

Comprehension question (easy): Was Morris on a nighttime run?

(15) a. Natasha saw an exhausted janitor on her way to a doctor's appointment. She started climbing the stairs to the doctor's office. {The janitor/He} accidentally knocked several cleaning bottles off his cart. Natasha turned back to help him pick them up.

b. Natasha saw an unattended mop on her way to a doctor's appointment.

She started climbing the stairs to the doctor's office. {The mop/It} was at the bottom of the stairwell. A janitor appeared and grabbed it.

Comprehension question (easy): Did Natasha get on an elevator?

(16) a. Owen saw a preschool teacher at the natural history museum. He snapped some photos of the fossil displays. {The preschool teacher/She} scolded two of her little students. They had tried to wander away from their group.

b. Owen saw a special exhibit at the natural history museum. He snapped some photos of the fossil displays. {The special exhibit/It} was only in town for a month. Owen loved the exhibit, although the museum was uncomfortably crowded.

Comprehension question (easy): Did Owen take photos of some paintings?

(17) a. Phyllis saw a doorman across the lobby of her friend's apartment building. She texted her friend that she was downstairs. {The doorman/He} glanced up at her only briefly. Apparently she looked harmless enough.

b. Phyllis saw a big mirror across the lobby of her friend's apartment building. She texted her friend that she was downstairs. {The mirror/It} made the lobby look much larger. Phyllis used it to inspect her makeup.

Comprehension question (easy): Did Phyllis call her friend?

(18) a. Quinton saw a wedding planner in a popular local bakery. He ordered some pastries and grabbed a table. {The wedding planner/She} was helping her clients explain their design ideas to the baker. It all sounded ridiculously overcomplicated to Quinton.

b. Quinton saw an impressive torte in a popular local bakery. He ordered some pastries and grabbed a table. {The torte/It} was decorated to look like a Monet painting. Quinton marveled at the baker's skill.

Comprehension question (easy): Did Quinton order some bagels?

(19) a. Renee saw an electrician at her sister's house the other day. She went into the living room. {The electrician/He} shut off the power. Renee switched her phone to battery-saver mode.

b. Renee saw a robotic vacuum at her sister's house the other day. She went into the living room. {The vacuum/It} followed her in. Renee hopped onto the couch to avoid it.

Comprehension question (easy): Was Renee at her mother-in-law's house?

(20) a. Stan saw a ballerina in the plaza near the dance theater. He crossed the plaza on his way to the train station. {The ballerina/She} was posing for a photo shoot. Stan made sure not to step into the shot by accident.

b. Stan saw a sculpture in the plaza near the dance theater. He crossed the plaza on his way to the train station. {The sculpture/It} was made by a venerated artist. Stan decided to look at it more closely on his return trip.

Comprehension question (easy): Was Stan on his way to the bus stop?

(21) a. Tanya saw a fisherman near the kelp beds during her diving session. She set up her diving buoys. {The fisherman/He} signaled to her. He showed her where his lines were so she wouldn't get tangled up.

Comprehension question (hard): Did Tanya set up her buoys before the fisherman signaled her?

b. Tanya saw a crab pot near the kelp beds during her diving session. She set up her diving buoys. {The crab pot/It} shifted on the seafloor. An octopus was trying to break into it.

Comprehension question (hard): Did Tanya set up her buoys before the crab pot moved?

(22) a. Victor saw a well-known actress at the table next to his in a quiet restaurant. He opened his menu. {The actress/She} let out a frustrated curse. The paparazzi had appeared outside the restaurant window.

Comprehension question (hard): Did Victor see someone famous?

b. Victor saw a fancy cocktail at the table next to his in a quiet restaurant. He opened his menu. {The cocktail/It} wasn't listed anywhere. Victor asked the waiter what it was.

Comprehension question (hard): Did Victor read the cocktail menu?

(23) a. Whitney saw a carpenter at the house next door to hers a couple days ago. She stretched out in a sunny spot in her backyard. {The carpenter/He} fired up his circular saw. He was cutting planks for a new deck.

Comprehension question (hard): Did the carpenter start his saw after Whitney stretched out?

b. Whitney saw a birdfeeder at the house next door to hers a couple days ago. She stretched out in a sunny spot in her backyard. {The birdfeeder/It} shook suddenly on its post. A squirrel had jumped onto it to steal the birdseed.

Comprehension question (hard): Did the squirrel jump on the birdfeeder after Whitney stretched out?

(24) a. Xavier saw a housekeeper in his rich in-laws' penthouse. He sprawled out on a huge sofa. {The housekeeper/She} gave him a judgmental look. She disapproved of people putting their feet on the furniture.

b. Xavier saw an oil portrait in his rich in-laws' penthouse. He sprawled out on a huge sofa. {The portrait/It} loomed over him somewhat creepily. It featured his husband's great-grandparents in stiff formal clothes.

Comprehension question (hard): Is Xavier married?

(25) a. Yvonne saw a flustered waiter at the café at the end of her block. She strolled by the café's outdoor patio. {The waiter/He} was trying to handle

a loud argument between two customers. Eventually the manager had to intervene.

b. Yvonne saw a large tent at the café at the end of her block. She strolled by the café's outdoor patio. {The tent/It} rustled in the breeze. A few customers were enjoying the shade it provided.

Comprehension question (hard): Is there a café near where Yvonne lives?

(26) a. Zachary saw a fortune teller on his bike ride home from work. He pedaled past a row of parked cars. {The fortune teller/She} shouted after him. She told him to beware of pigeons for the next week.

Comprehension question (hard): Did Zachary receive a warning?

b. Zachary saw a new shop on his bike ride home from work. He pedaled past a row of parked cars. {The shop/It} had a display of healing crystals in the window. Zachary made a mental note to check it out on the weekend.

Comprehension question (hard): Was Zachary interested in healing crystals?

(27) a. Alyssa saw a construction worker on the ferry the other day. She took her usual seat by the starboard windows. {The construction worker/He} relaced his boots and adjusted his vest. Alyssa guessed that he was headed to the new waterfront project.

b. Alyssa saw a backhoe on the ferry the other day. She took her usual seat by the starboard windows. {The backhoe/It} was parked on a flatbed truck. Alyssa guessed that it was headed to the new waterfront project.

Comprehension question (hard): Does Alyssa always sit in the same place on the ferry?

(28) a. Boris saw a receptionist on his way to his lawyer's office. He pressed the button for the elevator. {The receptionist/She} asked him where he was headed. He had neglected to sign in.

Comprehension question (hard): Did Boris press the elevator button before the receptionist talked to him?

b. Boris saw an indoor fountain on his way to his lawyer's office. He pressed the button for the elevator. {The fountain/It} dominated the foyer. Boris decided it was a gaudy waste of water.

Comprehension question (hard): Did Boris disapprove of the fountain?

(29) a. Carina saw a retired quarterback at the airport last week. She removed her shoes for the security screening. {The quarterback/He} paused just beyond the metal detector. A fan had recognized him and asked for a picture.

Comprehension question (hard): Was the quarterback ahead of Carina in the security line?

b. Carina saw a tacky suitcase at the airport last week. She removed her shoes for the security screening. {The suitcase/It} was on the belt in front of her shoes. It had a huge designer logo and a loud floral print on it.

Comprehension question (hard): Was the suitcase a carry-on?

(30) a. David saw a babysitter at the community pool on Thursday. He put on his goggles and swim cap. {The babysitter/She} shrieked with laughter. Her young charges were splashing her with water.

Comprehension question (hard): Did the babysitter laugh after David put on his goggles?

b. David saw a big umbrella at the community pool on Thursday. He put on his goggles and swim cap. {The umbrella/It} clattered to the ground. The wind had blown it over.

Comprehension question (hard): Did the umbrella fall after David put on his goggles?

(31) a. Elaine saw a barber on her way home from the grocery store. She switched her grocery bag from one hand to the other. {The barber/He} was chatting with a happy-looking customer. Elaine made a note to recommend him to her boyfriend.

b. Elaine saw a barbershop on her way home from the grocery store. She switched her grocery bag from one hand to the other. {The barbershop/It} was full of happy-looking customers. Elaine made a note to recommend it to her boyfriend.

Comprehension question (hard): Is Elaine married?

(32) a. Felix saw a softball player at the field over the weekend. He walked between the bleachers and the backstop. {The softball player/She} hit a ball launched by a pitching machine. The ball soared into right field.

Comprehension question (hard): Did Felix walk by the bleachers after the softball player hit the ball?

b. Felix saw a pitching machine at the field over the weekend. He walked between the bleachers and the backstop. {The pitching machine/It} launched a ball towards home plate. The ball thunked into the backstop.

Comprehension question (hard): Did Felix walk by the bleachers after the pitching machine launched the ball?

(33) a. Gloria saw a sportswriter at the basketball courts near her apartment. She took a few warm-up shots. {The sportswriter/He} flagged her down when she passed by. He wanted to ask her some questions for a story he was working on.

b. Gloria saw a ball cart at the basketball courts near her apartment. She took a few warm-up shots. {The ball cart/It} wobbled when she passed by. It had a broken wheel.

Comprehension question (hard): Was Gloria in the middle of a basketball game?

(34) a. Harry saw a nutritionist on his way to his training session. He started his stretching routine. {The nutritionist/She} was carrying an armload of supplements. She offered him a free consultation.

Comprehension question (hard): Did the nutritionist offer Harry supplements?

b. Harry saw a yoga ball on his way to his training session. He started his stretching routine. {The yoga ball/It} rolled into the backs of his knees. He kicked it back into place.

Comprehension question (hard): Did Harry pick up the yoga ball?

(35) a. Ilana saw a weightlifter at the gym two days ago. She hopped onto a treadmill. {The weightlifter/He} shouted in triumph. He had just lifted a personal best.

b. Ilana saw a weight rack at the gym two days ago. She hopped onto a treadmill. {The weight rack/It} rattled alarmingly. Someone had jostled it.

Comprehension question (hard): Did Ilana hop off the treadmill?

(36) a. Joel saw a figure skater at the rink on the east side of town. He laced up his skates. {The figure skater/She} landed an impressive jump. Joel gave her a cheer.

b. Joel saw a Zamboni at the rink on the east side of town. He laced up his skates. {The Zamboni/It} turned at the far end of the ice. Joel waited for it to finish smoothing the ice.

Comprehension question (hard): Was the skating rink on the west side of town?

(37) a. Karli saw a farmer in an apple orchard just outside town. She grabbed a crate to fill with apples. {The farmer/He} directed her to the pink lady apples. She made her way down the row he indicated.

Comprehension question (hard): Did Karli grab a crate after getting directions?

b. Karli saw a tractor in an apple orchard just outside town. She grabbed a crate to fill with apples. {The tractor/It} motored down an adjacent row of trees. Karli made her way towards the pink lady grove.

Comprehension question (hard): Did Karli grab a crate before spotting the tractor?

(38) a. Luke saw a gymnast at the physical therapy facility last night. He started preparing for his ice bath. {The gymnast/She} was using a foam roller.

She winced as it dug into her hamstrings.

Comprehension question (hard): Was Luke done with his ice bath when he saw the gymnast?

b. Luke saw a foam roller at the physical therapy facility last night. He started preparing for his ice bath. {The foam roller /It}toppled over with a thud. Luke tilted it back against the wall.

Comprehension question (hard): Was Luke done with his ice bath when he saw the foam roller?

(39) a. Monique saw a plumber at the hardware store downtown. She turned down the paint aisle. {The plumber/He} yelped in pain. A pipe had fallen on his foot.

Comprehension question (hard): Did Monique turn down the paint aisle after the plumber yelped?

b. Monique saw a forklift at the hardware store downtown. She turned down the paint aisle. {The forklift/It} beeped in warning. It was backing up with a heavy pallet.

Comprehension question (hard): Did Monique turn down the paint aisle after the forklift beeped?

(40) a. Nolan saw a hygienist at the dentist's office last week. He filled out his check-in form. {The hygienist/She} laughed loudly. She was chatting

with an elderly patient.

b. Nolan saw an x-ray machine at the dentist's office last week. He filled out his check-in form. {The x-ray machine/It} buzzed loudly. Another patient was getting their molars x-rayed.

Comprehension question (hard): Was Nolan at the dentist's office last month?

D.2 Experiment 3B: Event structure x RFC

(41) Jessie {saw/brushed} by an elderly man in front of a quaint farmhouse on her morning walk. She continued toward the end of the lane. {The man/He} fell over. Jessie circled back to investigate what had happened.

Comprehension question (easy): Was Jessie on a morning walk?

(42) Albert {saw/brushed} by a nurse at the bus stop on his way to work. He sat down to wait for the bus. {The nurse/She} dropped her phone. Albert picked it up for her.

Comprehension question (easy): Did Albert wait for the bus?

(43) Beth {saw/brushed} by a firefighter in the grocery store yesterday afternoon. She made her way down the dairy aisle. {The firefighter/He} put a gallon of milk in his cart. Beth stopped to peruse the yogurt selection.

Comprehension question (easy): Did Beth go down the dairy aisle?

(44) Curtis {saw/brushed} by a tall woman at the beach last weekend. He laid out his towel on the sand. {The woman/She} frowned at him. Curtis was too close to the beach volleyball court.

Comprehension question (easy): Did Curtis lay out a towel?

(45) Diana {saw/brushed} by a butcher at the farmer's market on Wednesday. She stopped at a couple of farm stalls. {The butcher/He} waved her over. Diana went to say hello.

Comprehension question (easy): Was Diana at the farmer's market?

(46) Edwin {saw/brushed} by a cheerleader in the hall in between classes. He shut his locker and headed to class. {The cheerleader/She} smiled at him. Edwin smiled awkwardly back.

Comprehension question (easy): Did Edwin head to class?

(47) Fatima {saw/brushed} by a surgeon on her visit to the hospital. She got into the elevator. {The surgeon/He} called after her. He asked her to hold the door.

Comprehension question (easy): Did Fatima visit the hospital?

(48) Gregory {saw/brushed} by a new librarian during his most recent weekly trip to the library. He turned in the books he'd finished. {The librarian/She} placed an armload of books on a cart. Gregory nodded at her as he passed.

Comprehension question (easy): Did Gregory turn in some library books?

(49) Holly {saw/brushed} by a young pilot at the airport the other day. She arrived at her gate and sat down. {The pilot/He} hurried past. Holly figured he was running late.

Comprehension question (easy): Was Holly at the airport?

(50) Isaac {saw/brushed} by a very busy waitress at his favorite diner. He ordered his usual lunch at the counter. {The waitress/She} hustled by with a huge tray. The diner was packed, but understaffed.

Comprehension question (easy): Did Isaac order lunch?

(51) Jade {saw/brushed} by a baseball player on the train a couple nights ago. She found a spot to stand near the doors. {The baseball player/He} chuckled at something on his phone. Jade only recognized him because her son was a baseball fanatic.

Comprehension question (easy): Was Jade in a taxi?

(52) Kelvin {saw/brushed} by a florist while he was walking home on Monday. He adjusted the strap of his messenger bag. {The florist/She} ran up to him. His gloves had fallen out of his bag.

Comprehension question (easy): Was Kelvin riding a motorcycle?

(53) Louise {saw/brushed} by a groundskeeper at the golf course last weekend. She

hefted her golf bag over her shoulder. {The groundskeeper/He} tossed some fertilizer over a brown patch of grass. Louise greeted him on her way to the next hole.

Comprehension question (easy): Was Louise at the basketball courts last weekend?

(54) Morris {saw/brushed} by a stylish woman during his morning run through the city. He paused at a stoplight, jogging in place. {The woman/She} hailed a cab. Morris wished his own fashion sense were as good as hers.

Comprehension question (easy): Was Morris on a nighttime run?

(55) Natasha {saw/brushed} by an exhausted janitor on her way to a doctor's appointment. She started climbing the stairs to the doctor's office. {The janitor/He} accidentally knocked several cleaning bottles off his cart. Natasha turned back to help him pick them up.

Comprehension question (easy): Did Natasha get on an elevator?

(56) Owen {saw/brushed} by a preschool teacher at the natural history museum. He snapped some photos of the fossil displays. {The preschool teacher/She} scolded two of her little students. They had tried to wander away from their group.

Comprehension question (easy): Did Owen take photos of some paintings?

(57) Phyllis {saw/brushed} by a doorman across the lobby of her friend's apartment building. She texted her friend that she was downstairs. {The doorman/He} glanced up at her only briefly. Apparently she looked harmless enough.

Comprehension question (easy): Did Phyllis call her friend?

(58) Quinton {saw/brushed} by a wedding planner in a popular local bakery. He ordered some pastries and grabbed a table. {The wedding planner/She} was helping her clients explain their design ideas to the baker. It sounded ridiculously overcomplicated to Quinton.

Comprehension question (easy): Did Quinton order some bagels?

(59) Renee {saw/brushed} by an electrician at her sister's house the other day. She went into the living room. {The electrician/He} shut off the power. Renee switched her phone to battery-saver mode.

Comprehension question (easy): Was Renee at her mother-in-law's house?

(60) Stan {saw/brushed} by a ballerina in the plaza near the dance theater. He crossed the plaza on his way to the train station. {The ballerina/She} was posing for a photo shoot. Stan made sure not to step into the shot by accident.

Comprehension question (easy): Was Stan on his way to the bus stop?

(61) Tanya {saw/brushed} by a fisherman near the kelp beds during her diving session. She set up her diving buoys. {The fisherman/He} signaled to her. He

showed her where his lines were so she wouldn't get tangled up.

Comprehension question (hard): Did Tanya set up her buoys before the fisherman signaled her?

(62) Victor {saw/brushed} by a well-known actress at the table next to his in a quiet restaurant. He opened his menu. {The actress/She} let out a frustrated curse. The paparazzi had appeared outside the restaurant window.

Comprehension question (hard): Did Victor see someone famous?

(63) Whitney {saw/brushed} by a carpenter at the house next door to hers a couple days ago. She stretched out in a sunny spot in her backyard. {The carpenter/He} fired up his circular saw. He was cutting planks for a new deck.

Comprehension question (hard): Did the carpenter start his saw after Whitney stretched out?

(64) Xavier {saw/brushed} by a housekeeper in his rich in-laws' penthouse. He sprawled out on a huge sofa. {The housekeeper/She} gave him a judgmental look. She disapproved of people putting their feet on the furniture.

Comprehension question (hard): Is Xavier married?

(65) Yvonne {saw/brushed} by a flustered waiter at the café at the end of her block. She strolled by the café's outdoor patio. {The waiter/He} was trying to handle a loud argument between two customers. Eventually the manager had to inter-

vene.

Comprehension question (hard): Is there a café near where Yvonne lives?

(66) Zachary {saw/brushed} by a fortune teller on his bike ride home from work. He pedaled past a row of parked cars. {The fortune teller/She} shouted after him. She told him to beware of pigeons for the next week.

Comprehension question (hard): Did Zachary receive a warning?

(67) Alyssa {saw/brushed} by a construction worker on the ferry the other day. She took her usual seat by the starboard windows. {The construction worker/He} relaced his boots and adjusted his vest. Alyssa guessed that he was headed to the new waterfront project.

Comprehension question (hard): Does Alyssa always sit in the same place on the ferry?

(68) Boris {saw/brushed} by a receptionist on his way to his lawyer's office. He pressed the button for the elevator. {The receptionist/She} asked him where he was headed. He had neglected to sign in.

Comprehension question (hard): Did Boris press the elevator button before the receptionist talked to him?

(69) Carina {saw/brushed} by a retired quarterback at the airport last week. She removed her shoes for the security screening. {The quarterback/He} paused

just beyond the metal detector. A fan had recognized him and asked for a picture.

Comprehension question (hard): Was the quarterback ahead of Carina in the security line?

(70) David {saw/brushed} by a babysitter at the community pool on Thursday. He put on his goggles and swim cap. {The babysitter/She} shrieked with laughter. Her young charges were splashing her with water.

Comprehension question (hard): Did the babysitter laugh after David put on his goggles?

(71) Elaine {saw/brushed} by a barber on her way home from the grocery store. She switched her grocery bag from one hand to the other. {The barber/He} was chatting with a happy-looking customer. Elaine made a note to recommend him to her boyfriend.

Comprehension question (hard): Is Elaine married?

(72) Felix {saw/brushed} by a softball player at the field over the weekend. He walked between the bleachers and the backstop. {The softball player/She} hit a ball launched by a pitching machine. The ball soared into right field.

Comprehension question (hard): Did Felix walk by the bleachers after the softball player hit the ball?

(73) Gloria {saw/brushed} by a sportswriter at the basketball courts near her apartment. She took a few warm-up shots. {The sportswriter/He} flagged her down when she passed by. He wanted to ask her some questions for a story he was working on.

Comprehension question (hard): Was Gloria in the middle of a basketball game?

(74) Harry {saw/brushed} by a nutritionist on his way to his training session. He started his stretching routine. {The nutritionist/She} was carrying an armload of supplements. She offered him a free consultation.

Comprehension question (hard): Did the nutritionist offer Harry supplements?

(75) Ilana {saw/brushed} by a weightlifter at the gym two days ago. She hopped onto a treadmill. {The weightlifter/He} shouted in triumph. He had just lifted a personal best.

Comprehension question (hard): Did Ilana hop off the treadmill?

(76) Joel {saw/brushed} by a figure skater at the rink on the east side of town. He laced up his skates. {The figure skater/She} landed an impressive jump. Joel gave her a cheer.

Comprehension question (hard): Was the skating rink on the west side of town?

(77) Karli {saw/brushed} by a farmer in an apple orchard just outside town. She

grabbed a crate to fill with apples. {The farmer/He} directed her to the pink lady apples. She made her way down the row he indicated.

Comprehension question (hard): Did Karli grab a crate after getting directions?

(78) Luke {saw/brushed} by a gymnast at the physical therapy facility last night. He started preparing for his ice bath. {The gymnast/She} was using a foam roller. She winced as it dug into her hamstrings.

Comprehension question (hard): Was Luke done with his ice bath when he saw the gymnast?

(79) Monique {saw/brushed} by a plumber at the hardware store downtown. She turned down the paint aisle. {The plumber/He} yelped in pain. A pipe had fallen on his foot.

Comprehension question (hard): Did Monique turn down the paint aisle after the plumber yelped?

(80) Nolan {saw/brushed} by a hygienist at the dentist's office last week. He filled out his check-in form. {The hygienist/She} laughed loudly. She was chatting with an elderly patient.

Comprehension question (hard): Was Nolan at the dentist's office last month?

D.3 Experiment 3: Fillers

(81) Wayne went to the farmers' market. He needed some inspiration to finish his weekly meal-planning. He checked out all the dairy stalls first. He got his favorite cheeses, and some he hadn't tried before.

Comprehension question (easy): Did Wayne buy cheese?

(82) Luis did a bunch of landscaping work in his garden. First, he installed a border of drought-resistant shrubs. Next, he built some vegetable beds. He moved on to flowers after that.

Comprehension question (easy): Does Luis have a garden?

(83) Heather had to deal with a couple of health problems last week. On Monday, she had an awful toothache. She went to the dentist to get it taken care of. A couple days later, another problem arose.

Comprehension question (easy): Did Heather have to go to the dentist?

(84) Joel was in a foul mood this morning. First, he discovered that his roommates had left a mess in the kitchen. Then he realized he was out of both coffee and cereal. He rushed to a cafe to get breakfast before work.

Comprehension question (easy): Did Joel's roommates do something inconsiderate?

(85) Last weekend, Mark realized that he had some unhealthy online habits. On

Saturday, he made several impulse buys. He somehow spent two hours that evening trying to take a perfect selfie. Sunday was no better.

Comprehension question (easy): Did Mark buy some stuff online?

(86) Mark argued heatedly with Renee about wedding seating arrangements. Ten minutes later, he accused her of ordering the wrong color of napkins. That set Renee off, because it was Mark who had ordered the napkins. She quickly left the house before she said something she'd regret.

Comprehension question (easy): Did Renee order the napkins?

(87) Lily crashed her bike into a low wall. A moment earlier, a spoke in the front wheel had snapped. Luckily, she wasn't badly injured, but her bike was pretty banged up. She had to take it to the repair shop in town.

Comprehension question (easy): Did Lily crash her car?

(88) Lizzie had a pet-related fiasco this past weekend. She was looking after her sister's cat, Nigel. Unfortunately, she and Nigel hated each other. Saturday passed in relative peace.

Comprehension question (easy): Was Nigel a dog?

(89) Arthur did some volunteer work on Saturday. He was part of an environmental justice organization. The group was fighting to get anti-fracking legislation passed. They needed to get several thousand signatures on a petition.

Comprehension question (easy): Did Arthur's organization have all the signatures they needed?

(90) Johnny and his son Kevin made a bit of a scene at the family reunion last weekend. Everyone had gathered at Johnny's house. Johnny and the other adults started off with snacks and drinks in the backyard. Meanwhile, Kevin and all his cousins played all over the house.

Comprehension question (easy): Were the adults hanging out in the kitchen?

(91) Isaiah was working on his laptop in the upstairs study. He turned on the air conditioner in the corner of the room. A moment later, a weird smell began to fill the room. He couldn't figure out what the smell was, or where it was coming from.

Comprehension question (hard): Does Isaiah live in a multi-story house?

(92) Elena is a professional soccer player. She criticized her team executives' lack of support for their own players. Twenty minutes later, the team's front office fined her $5000 for a supposed 'conduct violation'. They wouldn't say what exactly Elena had done wrong.

Comprehension question (hard): Did Elena get fined after criticizing her team's executives?

(93) Lorena was trying to win a pair of rare sneakers in an auction. She raised her

bid by a hundred dollars–four times the minimum bid increase. A moment earlier, another bidder had increased the bid by fifty dollars. Lorena refreshed the auction page every few minutes.

Comprehension question (hard): Did Lorena participate in an online auction?

(94) Emmy shut her dog, Walter, in the bedroom by himself. A half hour earlier, Walter had chewed up a bunch of shoes. In the bedroom, Walter gnawed on the furniture until Emmy let him out. She consulted a ton of dog-training websites to find a better way to handle him.

Comprehension question (hard): Did Walter chew up shoes before getting shut in the bedroom?

(95) This past weekend, Sean went to the gym a few miles from his house. He had just signed up for a membership as part of his new year's resolution. He struggled through some dumbbell exercises first. A moment later, he failed to jog just a quarter mile on the treadmill without stopping.

Comprehension question (hard): Did Sean find the dumbbell exercises difficult?

(96) Ryan spent some quality time with his dog Buster this weekend. He first took Buster on a nice long walk in the woods. Then they drove to the pet store. Ryan let Buster pick out a new chew toy.

Comprehension question (hard): Did Ryan and Buster go for walk after going to the pet store?

(97) Padma is a big science fiction fan. Last Wednesday, she devoted hours to her fan interests. First, she bought a limited edition figurine online. After that, she watched a livestreamed interview with her favorite sci-fi author.

Comprehension question (hard): Does Padma think figurines are a waste of money?

(98) Alicia had the starring role in a superhero movie until last Thursday. However, she abruptly quit the production. She posted a short statement about staying positive and moving forward. Ten minutes earlier, the online rumor mill had gone into overdrive about a possible scandal.

Comprehension question (hard): Was Alicia widely expected to quit the production?

(99) Gail started preparing a stew. She put in meat, vegetables, and broth. Then, she added some corn starch to the stew. Five minutes later, she realized that it was the wrong consistency.

Comprehension question (hard): Did Gail realize the stew's consistency was wrong before she added the corn starch?

(100) Neil went on a day trip to a small beach town. He hung out at the beach until

lunch time. First, he went for a swim. Then he did some shoreline fishing.

Comprehension question (hard): Did Neil go swimming after lunch?

(101) Jodie spent the day fishing on Saturday. At around 1 PM, she tied a new hook onto her line. Then she put fresh bait on the hook. A day later, she caught a large trout in the deepest part of the lake.

Comprehension question (easy): Did Jodie catch a trout?

(102) Dan threw out his back for the first time ever. Many hours earlier, he twisted awkwardly while trying to put on a sweater. That made his back hurt even more. He had to lie in bed trying not to move for the rest of the day.

Comprehension question (easy): Did Dan hurt his back?

(103) Phoebe sang a simple passage quite poorly during her singing lesson. Five hours later, she got really frustrated with how she sounded. She even thought about quitting singing altogether. That night, she could barely focus on her homework.

Comprehension question (easy): Did Phoebe feel frustrated?

(104) Allison attempted a DIY clock repair project last Thursday. First, she removed the mechanism from the wooden case. Several hours later, the mechanism began to emit a high-pitched sqeak. Patiently, Allison kept making little adjustments until the squeaking subsided.

Comprehension question (easy): Did Allison get the clock to stop squeaking?

(105) Diana got quite upset with herself on Friday. A day later, she dropped a mug and it shattered everywhere. That only made her feel worse. Eventually, though, she recovered some of the positivity she had started the day with.

Comprehension question (easy): Did Diana feel worse because she broke a mug?

(106) On Tuesday, Ernest finally got home well after dark. He collapsed into an overstuffed armchair by the fire. Half a day earlier, he had felt a wave of pain travel up his spine. He gingerly adjusted his posture and waited for the pain to pass.

Comprehension question (easy): Did Ernest get home early on Tuesday?

(107) Bonnie has a ton of plants in her apartment. Every evening she takes care of them. Last night, she did her rounds as usual. A day earlier, she had seen that her ficus looked droopy and pathetic.

Comprehension question (easy): Did Bonnie's ficus look healthy?

(108) On Thursday, Lottie had the house to herself overnight. She was a little nervous to be all alone at night. She flipped on the lights on the back patio. A couple days later, she heard a weird scratching noise in the yard.

Comprehension question (easy): Did Lottie turn on the lights in the front yard?

(109) Sam played Animal Crossing last night. First he planted some shrubs. Then, he began pulling all the weeds he could find. A day later, a villager offered to buy weeds from him for double the normal price.

Comprehension question (easy): Did the villager offer Sam the normal price for weeds?

(110) Last weekend, Travis dogsat for a friend of his. The dog, Rufus, was pretty high-maintenance. Travis gave an order to Rufus in an angry voice. Two days earlier, Rufus had run away from him at the dog park.

Comprehension question (easy): Was Rufus an easygoing dog?

(111) Kerry browsed in a mom-and-pop bookstore. Then, she went into a bakery on the corner. Ten minutes later, she became hungry enough to eat a bear. The bakery was warm, cozy, and full of delicious-looking treats.

Comprehension question (hard): Did Kerry get hungry after going into the bakery?

(112) Nellie had to do some emergency plumbing work this weekend. She shut off the water to the whole house. Five minutes later, she noticed water leaking from the base of the sink. Once the water was off, she mopped up the leak.

Comprehension question (hard): Did Nellie shut off the water before she noticed the leak?

(113) Eileen had an awful day at work yesterday. A coworker yelled at her for someone else's mistake. She returned from lunch much later than expected. Five minutes earlier, she had received a scolding email from her boss.

Comprehension question (hard): Was Eileen treated unfairly by her coworker?

(114) Rachel was reading her book by the fire. She noticed that the fire was almost out. A moment earlier, she had thrown a large amount of wood onto it. She watched the flames carefully until she was sure they wouldn't go out.

Comprehension question (hard): Did Rachel throw wood onto the fire before she noticed it was almost out?

(115) Last week, Kelly got herself into serious legal trouble. She worked at an upscale jeweler's shop downtown. She stuffed a handful of diamond bracelets into her bag. Fifteen minutes earlier, she had been fired and quickly escorted off the premises.

Comprehension question (hard): Did Kelly try to steal bracelets?

(116) Olive lost a coveted film role to a younger actor. A moment earlier, she had fired her longtime agent, a respected Hollywood power player. She began calling up old producer friends herself to see if they had anything for her. Unfortunately, she was too proud to accept the äging motherr̈oles they offered

up.

Comprehension question (hard): Did Olive fire her agent after she lost the role she wanted?

(117) Julie did some major repairs on her niece's teddy bear a few days ago. She restuffed the bear's arm with cotton batting. A moment earlier, a seam had burst at the end of its paw. Julie's progress on the repairs was slow, but she didn't get discouraged.

Comprehension question (hard): Did Julie restuff the bear's arm before a seam burst at the end of its paw?

(118) Katie put on rubber gloves before bleaching the sink. A moment later, she felt an itchy, burning sensation all over her hands. She hurriedly began washing her hands with soap and water. Then she put on some soothing hand lotion.

Comprehension question (hard): Did Katie's hands start itching before she put on the gloves?

(119) Helen had the flu last week. The first day, she felt overheated and dizzy all morning. In the afternoon, she started to feel even more feverish and uncomfortable. A moment later, she dragged yet another heavy fleece blanket over herself.

Comprehension question (hard): Were Helen's symptoms worse in the morning than in the afternoon?

(120) Pedro opened the windows on either side of his desk. Ten minutes later, the room became uncomfortably hot and muggy. Fidgety and increasingly sweaty, he kept trying to write. After an hour of making no progress, Pedro finally gave up and went for a walk to clear his head.

Comprehension question (hard): Did Pedro finish his work?

(121) Edith and Tony went on a date last Thursday. They went to Edith's favorite sushi bar first. Then they went bowling. They ended the night at the best ice cream shop in town.

Comprehension question (hard): Does Edith like sushi?

(122) Richard adopted a guinea pig over the weekend. He had never had a pet before. He spent hours researching guinea pig care beforehand. He was determined to be an excellent pet-parent.

Comprehension question (easy): Did Richard adopt a guinea pig a month ago?

(123) Maricela felt bored after work on Monday. She started a large jigsaw puzzle. A minute later, she fitted the last piece of it into place. She felt pretty accomplished.

Comprehension question (easy): Did Maricela start a small puzzle?

(124) Tamara saw a wrestler at the park the other day. She found a shady spot to sit. Then she took out her book and some snacks. He was taking selfies with a fan who had recognized him.

Comprehension question (hard): Did Tamara want to sit in the sun?

References

Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and philosophy*, *20*(1), 1–50.

Afantenos, S., & Asher, N. (2010). Testing sdrt's right frontier. *arXiv preprint arXiv:1006.5880*.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238.

Altshuler, D. (2016). *Events, states and times: An essay on narrative discourse in english*. Walter de Gruyter GmbH & Co KG.

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology*, *35*(3), 427–440.

Arnett, N., & Wagers, M. (2017). Subject encodings and retrieval interference. *Journal of Memory and Language*, *93*, 22–54.

Asher, N. (1996). Logical foundations of discourse structure and interpretation. In

Logic colloquium (pp. 1–45).

Asher, N. (2000). Discourse structure and the logic of conversation. In *Current research in the semantics pragmatics interface.*

Asher, N. (2008). Troubles on the right frontier. *Pragmatics and beyond new series*, *172*, 29.

Asher, N., & Lascarides, A. (1995). Lexical disambiguation in a discourse context. *Journal of semantics*, *12*(1), 69–108.

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

Asher, N., & Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, *115*(4), 591–610.

Austen, J. (2017). *Pride and prejudice* (Vol. 1). Artisan Shoppe.

Bailey, H. R., Kurby, C. A., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & cognition*, *45*(6), 940–955.

Banks, I. M. (2008). *Use of weapons*. Hachette UK.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Scheipl, F. (2012). Package 'lme4'. *CRAN. R Foundation for Statistical Computing, Vienna, Austria*.

Bates, D., Maechler, M., Bolker, B., Walker, S., et al. (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version*, *1*(7), 1–23.

Bates, E., MacWhinney, B., et al. (1989). Functionalism and the competition model. *The crosslinguistic study of sentence processing*, *3*, 73–112.

Bestgen, Y., & Costermans, J. (1994). Time, space, and action: Exploring the narrative structure and its linguistic marking. *Discourse Processes*, *17*(3), 421–446.

Bestgen, Y., & Vonk, W. (1995). The role of temporal segmentation markers in discourse processing. *Discourse Processes*, *19*(3), 385–406.

Bestgen, Y., & Vonk, W. (2000). Temporal adverbials as segmentation markers in discourse comprehension. *Journal of Memory and Language*, *42*(1), 74–87.

Black, A., Freeman, P., & Johnson-Laird, P. N. (1986). Plausibility and the comprehension of text. *British Journal of Psychology*, *77*(1), 51–62.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, *29*(4), 413–432.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of memory and language*, *34*(6), 774–806.

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, *80*(1), 1–28.

Caenepeel, M. (1991). Event structure vs. discourse structure. In *Proceedings of the* *dandi workshop on discourse coherence, edinburgh april.*

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1), 1–32.

Chiang, T. (1998). *Stories of your life and others*.

Comrie, B. (1985). *Tense* (Vol. 17). Cambridge University Press.

Costa, F., & Branco, A. (2012). Backshift and tense decomposition. In *Proceedings of the 19th international conference on head-driven phrase structure grammar, chungnam national university daejeon* (pp. 86–106).

Costermans, J., & Bestgen, Y. (1991). The role of temporal markers in the segmentation of narrative discourse. *Cahiers de psychologie cognitive*, *11*(3), 349–370.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the syntactic processor. *Natural language parsing: Psychological, computational and theoretical perspectives, Cambridge University Press, New York*.

Crinean, M., & Garnham, A. (2006). Implicit causality, implicit consequentiality and semantic roles. *Language and Cognitive Processes*, *21*(5), 636–648.

Dahlgren, K. (1988). *Naive semantics for natural language understanding*. Springer.

Davidse, K., & Vandelanotte, L. (2011). Tense use in direct and indirect speech in english. *Journal of Pragmatics*, *43*(1), 236–250.

Declerck, R. (1988). *Sequence of tenses in english.* Fac. van de Letteren en de Wijsbegeerte, Univ. Campus.

Declerck, R. (1995). Is there a relative past tense in english? *Lingua*, *97*(1), 1–36.

Declerck, R. (2005). The relation between temporal and modal uses of indicative verb forms. *Crosslinguistic Views on Tense, Aspect, and Modality, ed. by Bart Hollebrandse, Angeliek van Hout, and Co Vet*, *13*, 215–227.

Dickey, M. W. (2001). *The processing of tense: Psycholinguistic studies on the interpretation of tense and temporal relations* (Vol. 28). Springer Science & Business Media.

Dowty, D. R. (1979). Word meaning and montague grammar: The semantics of verbs and times in generative semantics and in montague's ptq (studies in linguistics and philosophy). dordrecht, holland: D. *Dordrecht: Reidel.*

Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and philosophy*, *9*(1), 37–61.

Drummond, A. (2016). *Ibex farm: Internet Based Experiments (Version 0.3.9)* [Computer program]. Retrieved from `http://spellout.net/ibexfarm/`

Duffy, S. A., Shinjo, M., & Myers, J. L. (1990). The effect of encoding task on memory for sentence pairs varying in causal relatedness. *Journal of Memory and Language*, *29*(1), 27–42.

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory?

Psychological science, *22*(2), 243–252.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current directions in psychological science*, *11*(1), 11–15.

Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of memory and language*, *25*(3), 348–368.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, *30*(6), 725–745.

Ferreira, F., & Henderson, J. M. (1998). Syntactic reanalysis, thematic processing, and sentence comprehension. In *Reanalysis in sentence processing* (pp. 73–100). Springer.

Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in english: A corpus of 300 verbs. *Behavior Research Methods*, *43*(1), 124–135.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior research methods*, *41*(1), 163–171.

Frazier, L. (1987). Syntactic processing: evidence from dutch. *Natural Language & Linguistic Theory*, *5*(4), 519–559.

Frazier, L. (1999). *On sentence interpretation* (Vol. 22). Springer Science & Business Media.

Frazier, L., & Clifton, C. (1996). *Construal*. Mit Press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, *19*(2), 101–131.

Garnham, A., Oakhill, J., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition*, *11*(1), 29–46.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Psychology Press.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive science*, *17*(3), 311–347.

Grice, H. (1975). Logic and conversation. *Syntax and Semantics*, *3*, 41-58.

Grosz, B., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*.

Haberlandt, K., Berian, C., & Sandson, J. (1980). The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior*, *19*(6), 635–650.

Hamm, F., & Bott, O. (2016). Tense and aspect. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Summer 2016 ed.). Metaphysics Research Lab, Stanford University. `https://plato.stanford.edu/archives/sum2016/entries/tense-aspect`

Hinrichs, E. (1986). Temporal anaphora in discourses of english. *Linguistics and philosophy*, *9*(1), 63–82.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive science*, *3*(1), 67–90.

Hobbs, J. R. (1985). On the coherence and structure of discourse.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes*, *26*(3), 376–405.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. *Meaning, form, and use in context: Linguistic applications*, *11*, 42.

Hsieh, Y., & Boland, J. E. (2015). Semantic support and parallel parsing in chinese. *Journal of psycholinguistic research*, *44*(3), 251–276.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Harvard University Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Kamp, H. (1981). A theory of truth and semantic representation. *Formal semantics-the essential readings*, 189–222.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*

(Vol. 42). Springer Science & Business Media.

Karimi, H., Diaz, M., & Ferreira, F. (2019). "a cruel king" is not the same as "a king who is cruel": Modifier position affects how words are encoded and retrieved from memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(11), 2010.

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

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

Kehler, A., & Rohde, H. (2017). Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes*, *54*(3), 219–238.

Köhne, J., & Demberg, V. (2013). The time-course of processing discourse connectives. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 35).

Komeda, H., & Kusumi, T. (2006). The effect of a protagonist's emotional shift on situation model construction. *Memory & cognition*, *34*(7), 1548–1556.

Kratzer, A. (1998). More structural analogies between pronouns and tenses. In *Semantics and linguistic theory* (Vol. 8, pp. 92–110).

Lascarides, A. (1992). *Knowledge, causality, and temporal representation.* Walter de Gruyter, Berlin/New York.

Lascarides, A., & Asher, N. (1993a). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and philosophy*, *16*(5), 437–493.

Lascarides, A., & Asher, N. (1993b). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and philosophy*, *16*(5), 437–493.

Lascarides, A., & Oberlander, J. (1993a). Temporal coherence and defeasible knowledge. *Theoretical linguistics*, *19*(1), 1–37.

Lascarides, A., & Oberlander, J. (1993b). Temporal connectives in a discourse context. In *Proceedings of the sixth conference on european chapter of the association for computational linguistics* (pp. 260–268).

Logačev, P., & Vasishth, S. (2016). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science*, *40*(2), 266–298.

Lorch, R. F., Lorch, E. P., & Matthews, P. D. (1985). On-line processing of the topic structure of a text. *Journal of memory and language*, *24*(3), 350–362.

Mandler, J. M. (1986). On the comprehension of temporal order. *Language and Cognitive Processes*, *1*(4), 309–320.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71.

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of memory and language*, *33*(1), 128–147.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference.

Computational linguistics, *14*(2), 15–28.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, *26*(4), 453–465.

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*(4), 569–587.

Noordman, L. G., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language*, *31*(5), 573–590.

Ogihara, T. (1995a). Double-access sentences and reference to states. *Natural Language Semantics*, *3*(2), 177–210.

Ogihara, T. (1995b). The semantics of tense in embedded clauses. *Linguistic inquiry*, 663–679.

Ohtsuka, K., & Brewer, W. F. (1992). Discourse organization in the comprehension of temporal order in narrative texts. *Discourse Processes*, *15*(3), 317–336.

Partee, B. H. (1973). Some structural analogies between tenses and pronouns in english. *The Journal of Philosophy*, *70*(18), 601–609.

Partee, B. H. (1984). Nominal and temporal anaphora. *Linguistics and philosophy*, *7*(3), 243–286.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of pragmatics*, *12*(5-6), 601–638.

Prasad, R. e. a. (2008). *Penn discourse treebank version 2.0* (Vol. LDC2008T05) (No.

Web Download). Philadelphia: Linguistic Data Consortium.

Prévot, L., & Vieu, L. (2008). The moving right frontier. *Pragmatics and beyond new series*, *172*, 53.

Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition*, *29*(8), 1073–1080.

Radvansky, G. A., & Copeland, D. E. (2004). Working memory span and situation model processing. *The American journal of psychology*, 191–213.

Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & cognition*, *34*(5), 1150–1156.

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 210.

Reichenbach, H. (1947). Elements of symbolic logic.

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition*, *28*(8), 1310–1320.

Rohde, H., & Horton, W. S. (2014). Anticipatory looks reveal expectations about discourse relations. *Cognition*, *133*(3), 667–691.

Rohde, H., Levy, R., & Kehler, A. (2011). Anticipating explanations in relative clause processing. *Cognition*, *118*(3), 339–358.

Rohde, H., Tyler, J., & Carlson, K. (2017). Form and function: Optional complementizers reduce causal inferences. *Glossa: a journal of general linguistics*, *2*(1).

Silverstein, M. (1976). *Hierarchy of features and ergativity. grammatical categories in australian languages.* RMW Dixond). Pp.

Simner, J., & Pickering, M. J. (2005). Planning causes and consequences in discourse. *Journal of Memory and Language*, *52*(2), 226–239.

Smith-Stark, T. C. (1974). The plurality split. In *Chicago linguistic society* (Vol. 10, pp. 657–672).

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological science*, *12*(2), 153–156.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216.

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 431.

Team, R. C. (2013). R: A language and environment for statistical computing.

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & cognition*, *34*(1), 78–89.

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology Section A*, *50*(3), 481–497.

Traxler, M. J., Pickering, M. J., & Clifton Jr, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*(4), 558–592.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes*, *6*(4), 303–338.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension.

van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*(2), 225–258.

Webber, B. L. (1988). Tense as discourse anaphor. *Computational Linguistics*, *14*(2), 61–73.

Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., ... Wasow, T. (2004). Animacy encoding in english: Why and how. In *Proceedings of the workshop on discourse annotation* (pp. 118–125).

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, memory, and cognition*, *22*(5), 1196.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological science*, *6*(5), 292–297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of experimental psychology: Learning, memory, and cognition*, *21*(2), 386.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS one*, *7*(12), e51382.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, *123*(2), 162.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological science*, *13*(2), 168–171.