UNIVERSITY OF CALIFORNIA
SANTA CRUZ

THE SOUND OF SILENCE: INVESTIGATIONS OF IMPLICIT PROSODY

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

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Nicholas J. Van Handel

September 2022

The Dissertation of Nicholas J. Van Handel
is approved:

Assistant Professor Amanda Rysling, chair

Professor Matthew Wagers

Associate Professor Ryan Bennett

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

The Sound of Silence: Investigations of Implicit Prosody

by

Nicholas J. Van Handel

This dissertation is about implicit prosody, the prosodic structure that readers assign during silent reading. The dissertation has several goals: determining which reading tasks are appropriate for studying implicit prosody, establishing how grammatical principles could guide incremental assignment of prosodic structure, and investigating how implicit prosody interacts with other properties such as focus in order to influence syntactic parsing and interpretation. On the methodological side, this dissertation demonstrates that the Maze task is suitable for studying implicit prosody by replicating several major findings on metrical and phrasal prosody in both the Maze task and in self-paced reading. In addition to showing that the Maze is sensitive to implicit prosody, the methodological comparison confirms previously reported advantages of the Maze over self-paced reading, such as more localized and larger effects. On the theoretical side, the dissertation lays the groundwork for developing an incremental model of prosodic parsing. I provide an overview of the major grammatical constraints that govern the syntax-prosody interface, drawing on work from the theoretical phonology literature. I discuss how and when these grammatical constraints, which are typically invoked to model the final phrasing

for a complete sentence structure, could be deployed by an incremental parser that assigns a prosodic structure word-by-word. Using a toy model of an incremental parser, I also show how the parser's first pass implicit prosody may differ from the final prosody, arguing that future work in this area should more closely consider these potential differences. The final set of experiments investigates both the timing of implicit prosodic assignment and how prosodic structure and information structure affect attachment decisions. Based on the results of these experiments, I propose the Visibility First Hypothesis, according to which attachment decisions are determined primarily by prosodic visibility, while other factors such as focus only exert an influence when two potential attachment sites are equally visible. I then outline several experiments to test the Visibility First Hypothesis in future work.

Acknowledgments

First and foremost, I would like to thank Amanda Rysling for being a generous, supportive, and patient advisor. Thank you for making time for last-minute meetings and reading countless drafts, for helping me work through my self-doubt, for seeing the potential in my ideas while challenging me to push my work further, for sharing your intimidatingly vast knowledge of phonetics, phonology, and psycholinguistics, and for teaching me how to channel spite and skepticism into something productive. These were the best days of my flerm. Thank you to Matt Wagers for introducing me to sentence processing in my first year of grad school, for providing encouragement when my experiments did not turn out as expected, and for being a supportive mentor on various projects. Thanks to Ryan Bennett for teaching me so much about phonetics and phonology, for providing invaluable advice on many projects, and for letting me spend time with Turnip and Digby.

I would also like to thank Junko Ito and Armin Mester for their mentorship and for the many fun and productive SPOT meetings, Grant McGuire for sharing his knowledge of cool phonetics studies, Pranav Anand for helping me out with semantics, Adrian Brasoveanu for his assistance with stats, and Ivy Sichel for her guidance during my first time teaching a class. Thanks to Kristin Van Engen, Brett Hyde, and Jungmin Kang, who introduced me to linguistics and helped me get involved in research as an undergraduate. Thanks also to Shayne Sloggett for feedback on various ideas, technical assistance with Ibex, and help with stats.

Many thanks to Netta Ben-Meir, whose friendship got me through grad school. Thanks for the hours we spent working together on problem sets and experimental items, for having the best recommendations for TV shows (both good and bad), and for always being there. Thanks also to my Curve family, Jed Pizarro-Guevara and Kelsey Sasaki. Jed, thanks for sharing your expertise in stats and experimental design, for singing everything from “Truly Madly Deeply” to “Blessed Be the Lord My Rock” with me, and for always being willing to drive us to San Jose for food. Kelsey, thanks for putting up with us and for being one of the most generous people I know.

Thanks also to Stephanie Rich and Lalitha Balachandran for being the best keepers and for collaborating with me on the experiments in Chapter 4. I learned so much from talking about psycholinguistics and prosody with you both. Stephanie, thanks for the sprinkles, and Lalitha, thanks for Taco Bell.

Thanks to Vishal Arvindam for being an amazing TA, to Jenny Bellik and Nick Kalivoda for sharing their syntax-prosody expertise, to Richard Bibbs for co-developing California Englishes, to Dan Brodkin and Ben Eischens for being great collaborators, to Ben again, for finally learning my name, to Morwenna Hoeks for always being willing to answer my questions about semantics, to Max Kaplan for sharing pictures of Cricket, and to Matthew Kogan for all of his help as my RA.

And thanks to the many other grad students who made the UCSC linguistics department such a great place to spend six years: Andrew Angeles, Jeremie Beauchamp, Dhyana Buckley, Mansi Desai, Jack Duff, Steven Foley, Andrew Hedding, Lisa Hofmann, Kelsey Kraus, Margaret Kroll, Maho Morimoto, Tom Roberts, Deniz Rudin, Jake Vincent, Maya Wax Cavallaro, Niko Webster, Anissa Zaitsu, and Erik Zyman.

Thanks also to Meg, Sam, Caitlin, Lisa, Lora, Kevin, Sarah, Tom, Lucy, Becca, and Bryanna for their friendship and for helping me stay sane whenever I needed a break from linguistics. Finally, thanks to my mom, my dad, and Kristen for all of their support. I couldn't have done this without you.

Chapter 1

Introduction

1.1 Prosody and language comprehension

The prosodic structure of sentences - their rhythm, phrasing, and intonational contour - influences the way perceivers comprehend sentences. Consider the NP/Z ambiguity in (1): the subordinate clause contains an optionally transitive verb (*leave*) followed by an ambiguous NP (*the house*) that could serve as the direct object of the verb *leaves* in the NP parse (1a) or as the subject of the main clause in the Z parse (1b). The principle of Late Closure states that the parser prefers to attach incoming material to the phrase that is currently being processed; accordingly, the parser will analyze the ambiguous NP as the object of the subordinate clause verb and expects the continuation in (1a) (Frazier, 1979; Frazier & Rayner, 1982). Processing difficulty arises when subsequent material disconfirms this parse: in (1b), the verb *is* indicates that the NP should have been analyzed as the subject of the main clause, resulting in the garden path effect.

(1) When Roger leaves the house...

  a. ...it's dark. (NP parse)

  b. ...is dark. (Z parse)

This syntactic difference is accompanied by a prosodic difference. In (1a), it is most natural to place a prosodic boundary after the NP *the house* while in (1b) it is most natural to place a boundary after the verb *leaves*; in both cases, this boundary coincides with the edge of the subordinate clause. Kjelgaard and Speer (1999) leveraged the prosodic differences between these sentences to ask whether prosodic structure influences syntactic parsing. In a series of judgment and cross-modal priming tasks, they played listeners recordings of NP and Z sentences and manipulated whether the sentence was presented with cooperating prosody (with the canonical boundary for the target structure), conflicting prosody (with the canonical boundary for the competing structure), or baseline prosody (compatible with either structure). They found that the typical penalty for Z sentences disappeared when presented with cooperating prosody, while the preference for the NP parse persisted with conflicting and baseline prosody. Listeners avoided a garden path when the prosodic boundary coincided with the target syntactic boundary, and Kjelgaard and Speer concluded that prosody has an immediate effect on the assignment of syntactic structure.

While Kjelgaard and Speer (1999) focused on the NP/Z ambiguity, consideration of prosody has shaped our understanding of other well-known sentence processing effects, including attachment height preferences, such that the degree and location of a boundary influences whether a phrase is attached high or low (Carlson, Clifton, & Frazier, 2001; Fodor, 1998, 2002b, i.a.), and local coherence effects, such that the parser is more likely to analyze a complement clause as a standalone sentence when it exhausts a prosodic phrase (Frazier, Clifton, Carlson, & Harris, 2014). The pervasive role of prosody is undoubtedly a consequence of the fact that prosody reflects syntax and information structure; see Cutler, Dahan, and Van Donselaar (1997) and Wagner and Watson (2010) for overviews of prosody in language comprehension.

1.1.1 Prosodic boundaries and attachment

A longstanding question in prosodic processing concerns how prosodic boundaries influence syntactic attachment decisions. Previous research has found that listeners are less likely to attach incoming material to a word that is followed by a prosodic boundary (Bishop, Chong, & Jun, 2015; Carlson et al., 2001; Fromont, Soto-Faraco, & Biau, 2017; Maynell, 1999; Schafer, 1997). The example in (2) from Carlson and Potter (2022) illustrates this point. This string is globally ambiguous, because the prepositional phrase *on Monday* could modify either *heard* or *had arrived*. Listeners' preferred interpretation is influenced by prosody: in (2a), a boundary intervenes between *heard* and the rest of the sentence, such that *arrived* and *on Monday* are grouped together prosodically. In this case, listeners are more likely to interpret the PP as modifying the embedded verb. In (2b), however, a boundary intervenes between the embedded verb and the PP, and listeners are more likely to report that the PP attaches to the matrix verb.

(2) a. Sam heard % that Bill had arrived on Monday.

b. Sam heard that Bill had arrived % on Monday.

Various proposals have been put forth to explain the connection between prosodic boundaries and non-local attachment. I will use the term Repellent Boundaries Hypothesis (RBH), stated in (3), to refer to the idea that a prosodic boundary encourages non-local attachment because it signals that the word preceding the boundary likely belongs to a different syntactic constituent than the incoming material.

(3) **Repellent Boundaries Hypothesis**: The parser is less likely to target a potential attachment site when it is immediately followed by a prosodic boundary, because boundaries can signal the end of a syntactic constituent, which decreases the likelihood that a word preceding a boundary would be syntactically grouped with incoming material.

The RBH encompasses several previous proposals about the role of prosodic boundaries in sentence processing. For instance, the Anti-Attachment Hypothesis of Watson and Gibson (2004a, 2005) states that listeners prefer not to attach an incoming word to a head that is immediately followed by an intonational phrase boundary. In Schafer's (1997) Prosodic Visibility Hypothesis, the ease of attachment to any given site is determined by the visibility of that attachment site, where visibility is dependent on the prosodic structure. An attachment site is most visible when it is in the prosodic phrase that is currently being processed but is less visible if it is outside of that phrase, i.e., if a boundary intervenes between the attachment site and the material currently being processed. Visibility is gradient, such that attachment sites are less visible as more prosodic boundaries intervene between the attachment site and the material being processed. The difference between the Anti-Attachment Hypothesis and the Prosodic Visibility Hypothesis is that the former only makes reference to the single boundary after the potential attachment site, whereas the latter depends on the overall prosodic structure of the sentence, because the visibility of material in previously processed prosodic phrases decreases as each successive boundary is encountered.

1.1.2 Accents, focus, and attachment

Boundaries are not the only prosodic feature that influences attachment: potential attachment sites that bear a pitch accent are also more likely to attract incoming material. In a series of experiments, Schafer, Carter, Clifton, and Frazier (1996) had participants listen to globally ambiguous sentences like (4) in which a relative clause could attach to either of the two nouns in a complex NP structure. The experiment manipulated whether there was an accent on either N1 (4a) or on N2 (4b). Participants answered questions probing whether they interpreted the relative clause as modifying N1 or N2.

(4) a. The tourists admired the MUSEUM of the city that they visited again in August.

b. The tourists admired the museum of the CITY that they visited again in August.

Schafer et al. (1996) found that participants provided more high attachment responses when N1 was accented, suggesting that information structure, as conveyed by the pitch accent, influenced ambiguity resolution. Based on these results, they proposed the Focus Attraction Hypothesis, stated in (5).

(5) **Focus Attraction Hypothesis**: It is more likely that a phrase that is neither a complement nor syntactically obligatory will be taken to modify a phrase

P if P is focused than if it is not, grammatical and pragmatic constraints permitting.

1.1.3 The interplay of boundaries and accents

Several studies have attempted to tease apart the relative contributions of focus (accents) and prosodic boundaries. Carlson and Tyler (2018) tested the Boundary-Dependent Hypothesis, which states that accents only affect attachment if the prosodic boundaries already favor an interpretation, because prosodic phrasing is more closely related to syntactic grouping than accents. They tested this by varying accent position and the presence of a boundary in sentences in which an adverbial phrase such as *last week* could attach to either the matrix or complement clause, as in (6); subsequent experiments tested similar manipulations in other structures. Under the BDH, they expected that an accent on the matrix verb would lead to more high attachment responses in (6a) compared to (6b), because the boundary already favored high attachment. In contrast, they predicted no impact of an accent on attachment responses in (6c) and (6d), because an accent on the matrix verb would be insufficient to overcome the structure's strong low attachment bias in the absence of a boundary also favoring high attachment. However, they failed to find an interaction of accent placement and boundary presence; the accent affected responses regardless of the presence of a boundary. They concluded that the effect of accents on attachment is independent of boundaries.

(6) a. John CLAIMED that Mary arrived IPh last week.

b. John claimed that Mary ARRIVED IPh last week.

c. John CLAIMED that Mary arrived last week.

d. John claimed that MARY arrived last week.

However, Lee and Garnsey (2012) found evidence that boundaries and accents may interact. In their study, they used sentences like (7) in which a relative clause could attach to either of two nouns in a complex NP structure. They manipulated whether there was an accent on N1 or N2 and whether an intonational phrase boundary appeared after N1 or after N2. An example item is shown in (7). The single slash indicates intermediate (weak) phrase boundaries, while the double slash indicates intonational (strong) phrase boundaries; accented words were always followed by either an intermediate or an intonational phrase boundary.

(7) a. The reporter interviewed the **DAUGHTERS** / of the hostage // who

b. The reporter interviewed the daughters of the **HOSTAGE** // who

c. The reporter interviewed the **DAUGHTERS** // of the hostage who

d. The reporter interviewed the daughters // of the **HOSTAGE** / who

They found a difference in responses in the Late Boundary conditions (7a,b), such that more high attachment responses were provided when the accent was on

N1. However, there was no effect of accent placement in the Early Boundary conditions (7c,d), when the intonational phrase boundary came after N1. They reasoned that this was evidence against the Visiblity Hypothesis, which they claim would have predicted a difference in attachment preferences between (7c) and (7d), with more low attachment responses in (7c), because the intermediate phrase boundary in (7d) rendered N2 less visible. Instead, they suggest that this is consistent with the Rational Speaker Hypothesis (Carlson et al., 2001; Clifton, Carlson, & Frazier, 2006), which states that listeners are sensitive to the reasons why speakers produced boundaries and interpret them accordingly. For instance, listeners who hear a particularly strong prosodic boundary after a very long constituent may not represent this boundary as particularly strong, because the reason for the boundary's strength may be due to the production pressures arising from the challenges of producing a long constituent, rather than reflecting the syntactic attachment intended by the speaker. Under the RSH, the intermediate phrase boundary after N2 may not have been "informative about syntax because there was a larger intonational phrase boundary after the high noun earlier in the utterance." However, if listeners discounted the lower intermediate phrase boundary, then the account cannot explain why the presence of an accent on N2 in (7d) did not lead to more low attachment interpretations in that condition relative to (7c), as opposed to the reported finding of an equal rate of high attachment responses.

Other interpretations of Lee and Garnsey's (2012) findings are possible. Carlson and Potter (2022) suggest that this finding could be an accidental null or the result of boundaries and accents canceling each other out. I propose an alternative: the Visibility First Hypothesis, which is stated in (8). Under this view, the parser first prioritizes prosodic visibility when attaching incoming material: attachment to the more visible attachment site is preferred. However, when two attachment sites are in the same phrase (i.e., they are equally visible), then the parser's choices can be affected by other considerations like focus and pitch accents.

(8) **Visibility First Hypothesis**: The parser preferentially attaches incoming material to the most visible potential attachment site, because prosodic visibility serves as a proxy for syntactic grouping. The parser weighs additional evidence, such as focus, only when two attachment sites are equally visible, because in this case visibility is insufficient for determining the most likely syntactic grouping.

This account can explain why boundaries and accents interacted in Lee and Garnsey's findings, provided that the effect of weaker intermediate phrase boundaries on visibility is negligible.[1] In the comparison of (7a) and (7b), both nouns are equally visible, because they are both separated from the RC by the strong in-

[1]This account is also consistent with a world in which intermediate phrase boundaries modulate visibility, but to a smaller extent than intonational phrase boundaries, such that focus could still attract attachment to a higher site even when there is an intervening intermediate phrase boundary. In this case, the VFH would need to specify that focus only modulates attachment preferences when both potential sites are in the same *intonational* phrase.

tonational phrase boundary. Accents now play a role and focus attracts attachment to N1 in (7a), leading to more high attachment in this condition. However, in (7c) and (7d), N2 is more visible than N1, because N1 is separated from the RC by an intonational phrase boundary. Since focus does not affect attachment when there is a major difference in visibility, the pitch accent on N2 in (7d) has a negligible effect, leading to comparable rates of high attachment interpretations in these two conditions.

At first glance, this account may seem at odds with Carlson and Tyler (2018), who rejected the idea that accents depend on boundaries. However, their conclusion was based on experiments that only tested the presence or absence of an intonational phrase boundary after the second attachment site. That is, their experiments always involved sentences in which both attachment sites were equally visible. Thus, the Visibility First Hypothesis is still compatible with their findings, as the VFH would also predict that accents should have an effect whenever both attachment sites are equally visible.

1.2 Implicit prosody: An overview

I have motivated the Visibility First Hypothesis on the basis of listening studies. However, the role of prosody is not limited to spoken language: readers generate a prosodic representation during silent reading, and this *implicit prosody* has consequences for syntactic parsing (Bader, 1998; Fodor, 1998, 2002a, 2002b). Understanding the role of implicit prosody is crucial for any theory of language comprehension, especially when one considers that psycholinguistic data often comes from studies that use written stimuli. Yet studying implicit prosody is difficult, because there are often many ways to prosodify any given string, and it is impossible to directly study the content of the phonological representations generated during reading to determine which prosodic structure readers assigned.

The role of implicit prosody in sentence processing has received increased attention in the past few decades. One of the first attempts to take seriously the idea that prosody generated during silent reading impacts syntactic parsing came from Fodor (1998, 2002a, 2002b), who proposed the Implicit Prosody Hypothesis:

(9) **Implicit Prosody Hypothesis (IPH)**: In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction.

Fodor appealed to the IPH in order to explain cross-linguistic differences in attachment site preferences in sentences like (10) and (11), in which the relative clause *who was on the balcony* could modify either *the servant* (high attachment) or *the actress* (low attachment). In English, comprehenders prefer the more local attachment site: the relative clause is interpreted as describing the actress. The bias for local attachment was argued to derive from Late Closure, a purportedly universal parsing principle. However, Spanish comprehenders show a preference in the opposite direction: in (11), there is a bias to interpret the relative clause as attaching to the non-local noun phrase, *la criada*.

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

(11) Alguien disparó contra la criada de la actriz que estaba en el balcón.

Fodor proposed that implicit prosody is responsible for this cross-linguistic variation. A prosodic break before a relative clause favors high attachment, but differences in languages' prosodic grammars affect the likelihood that a comprehender will place a break before the relative clause, which results in cross-linguistic differences in attachment preferences. French, which patterns like Spanish, favors a prosodic break before long relative clauses; this break is then interpreted as evidence of high attachment. English, in contrast, does not place a prosodic break before a relative clause by default, and the absence of a break leads the parser to adopt a low attachment analysis. Crucially, this reasoning relies on the assumption that the parser is sensitive to implicit prosody.

In another appeal to implicit prosody, Bader (1998) proposed the Prosodic Constraint on Reanalysis (12). According to the PCR, garden path sentences are more difficult to process when they require both syntactic and prosodic reanalysis than when they require only syntactic reanalysis, because prosodic reanalysis affects the reader's inner voice, which causes the reader to become aware of the processing disruption caused by the garden path.

(12) **Prosodic Constraint on Reanalysis (PCR):** Revising a syntactic structure is difficult if it necessitates a concomitant reanalysis of the associated prosodic structure.

The PCR explains why the NP/Z ambiguity in (13) is palpably more difficult than the NP/S ambiguity in (14). In both cases, there is a temporarily ambiguous NP that could be analyzed as either a direct object of the preceding verb or as the main subject of a new clause: *the little boy* in (13) and *the answer* in (14). The parser initially prefers the direct object analysis; in the (b) examples, the presence of a verb following the NP disconfirms this initial parse, and the parser must reanalyze the NP as the subject of a new clause. In (13), this reanalysis also requires a change to the prosodic structure: the initial parse places an intonational phrase boundary after *the little boy*, as in (13b.i) but in the target structure the boundary occurs earlier, after the verb *help*, as in (13b.ii). The parser must remove the implicit prosodic boundary it placed after *the little boy* and generate a new boundary in the correct location, and this prosodic reanalysis causes the comprehender to become aware of the garden path. In (14), both parses consist of a single intonational phrase; no major prosodic reanalysis is necessary when the direct object parse is disconfirmed in (14b), and the processing penalty is less severe.

(13) a. In order to help *the little boy* Jill put down the package she was carrying.

i. ($_\iota$ In order to help *the little boy*) ($_\iota$ Jill put down the package she was carrying.)

b. In order to help *the little boy* put down the package he was carrying.

i. *($_\iota$ In order to help *the little boy*) ($_\iota$ put down the package he was carrying.)

ii. ($_\iota$ In order to help) ($_\iota$ *the little boy* put down the package he was carrying.)

(14) a. ($_\iota$ Peter knew *the answer* immediately.)

b. ($_\iota$ Peter knew *the answer* would be false.)

In a self-paced reading experiment in German, Bader provided further support for the PCR by manipulating whether a garden path sentence required only syntactic reanalysis or both syntactic and prosodic reanalysis. He leveraged a *dative / possessor* ambiguity similar to (15), in which the pronoun *her* could function either as an indirect object of the verb *gave* (15a) or as the possessor of the noun *money* (15b). The presence of syntactic reanalysis was manipulated by using clause-final verbs that did or did not subcategorize for a dative object, while prosodic reanalysis was manipulated via the presence of the focus particle *ihr* requiring stress on the pronoun in the dative structure. He showed that processing difficulty arose only when both types of reanalysis were necessary, providing evidence that implicit prosody generated during silent reading impacts syntactic parsing and showing that explicit consideration of implicit prosody can shed light on long-standing issues in sentence processing, like reanalysis.

(15) a. Mary said that someone gave her money on her birthday.

b. Mary said that someone gave her money to Peter.

While the findings reviewed here have focused primarily on the effects of implicit prosodic boundaries, readers have been shown to be sensitive to other aspects of prosody such as metrical structure and pitch accents (Bader, 1998; Breen & Clifton, 2011; Breen, Fitzroy, & Oraa Ali, 2019; Kentner, 2012; Speer & Foltz, 2015; Stolterfoht, Friederici, Alter, & Steube, 2007); see Breen (2014) for a detailed overview. These findings make it clear that psycholinguistics must consider the role of implicit prosody in sentence processing: there is ample evidence that readers generate an implicit prosodic representation during silent reading and that this representation has consequences for our understanding of fundamental aspects of sentence processing like attachment and reanalysis.

1.3 Default prosody and incrementality

Fodor's (2002a, 2002b) Implicit Prosody Hypothesis claims that readers project a "default prosodic contour" onto written sentences, and this default prosody can then influence the syntactic analysis that readers adopt. However, this notion of "default prosody" is underspecified: there are often multiple licit ways to phrase any given string of words, and many factors condition phrasing, including information structure, speech rate, and individual variation. For implicit prosody, overt productions are often taken to reflect the default prosody associated with any given sentence (Fodor, 2002a, 2002b; Hirose, 2003; Hwang & Schafer, 2009); however, this still does not address the issue that multiple phrasings are typically available. Moreover, previous work has suggested that overt productions may not straightforwardly reflect the default prosody assigned during silent reading (Jun, 2010).

The idea that there is a default prosody is not unique to psycholinguistics: parallels exist in the theoretical phonology literature. In this area, it is typically assumed, implicitly or explicitly, that some phrasings are more marked than others, and that the aim of the theoretical account should be to model the most unmarked, "all else being equal" prosody. For instance, in his analysis of phonological phrasing in Italian, Ghini (1993) states that there is a preference for phrases containing two words, yet points out that "slowing down the speech rate to an adagio" can lead to phrases with a single word, although this is "highly marked" and the "complications exceed the scope of the paper." Even within the unmarked "moderato" speech rate, Ghini shows that there are often multiple possible unmarked phrasings. Similarly, in an analysis of Japanese, Ito and Mester (2013) acknowledge that "It is always possible, in a diligent pronunciation, to parse each word as a separate ϕ....But in usual speech, this is not what happens." In response to this variability, phonologists typically set aside a particular set of data to model, often those phrasings that are most frequent ("usual speech"), or at a neutral speech rate. The important point is that there is a parallel between the notion of default implicit prosody and the work of theoretical phonologists: even if the criteria for identifying this default prosody are not always clear, researchers in both areas have recognized that there is usually some phrasing that is the most neutral and/or natural, and have centered this phrasing in their work.

The incremental assignment of implicit prosodic structure has also been under-examined in previous work, and this presents another issue for the appeal to default prosody: the prosody assigned during first pass parsing could differ from the default prosody that is ultimately assigned once the parser has seen the entire construction. The lack of a theory of how implicit prosody gets assigned incrementally makes it difficult to identify these potential discrepancies between first pass parsing and default prosody, and raises many questions about when and how the parser uses different sources of evidence to build structure.

Here, I illustrate the challenge posed by not considering incrementality using

Fodor's account of RC length effects on attachment. There is a cross-linguistic tendency for ambiguous RCs in sentences like (16) to attach low (to *the senior*) when the RC is short, as in (16a), and to attach high (to *the friend*) when the RC is long, as in (16b). Again, Fodor (2002a, 2002b) appeals to implicit prosody: readers are more likely to place an implicit prosodic boundary before a long RC, and this boundary encourages a high attachment parse.

(16) The professor saw the friend of the senior...

a. ...who studied.

b. ...who studied for hours in the library.

While the length-based explanation is plausible, implementing this principle is not straightforward when one considers how the sentence in (16) is processed incrementally. When the parser first encounters the relative pronoun *who*, it has no way of knowing whether the RC will be short (16a) or long (16b). Instead, the parser must wait until it has already processed (a significant portion of) the RC to know whether the RC is long enough to motivate a prosodic boundary.

In principle, there are various ways the parser could deal with the absence of information about relative clause length at the moment that the parser first enters the relative clause. The first question is whether the parser would make any commitments about the presence or absence of a prosodic boundary between the complex NP and the RC before determining the length of the RC. One possibility is that the parser would first decide whether to posit a boundary on the basis of the preceding material, e.g., the length of the preceding complex NP. Another possibility is that the parser would leave the prosodic structure in the pre-RC region underspecified until it has ascertained additional information about the length of the relative clause.

Next, there is a question of when the parser would have sufficient information to decide to assign (or modify) prosodic structure on the basis of RC length. For instance, the parser could wait until it has reached the right edge of the relative clause, at which point it would determine whether the RC was short, supporting the absence of a pre-RC boundary, or long, motivating the presence of a boundary. Alternatively, the parser may not have to wait until the end of the RC to make a decision about length: perhaps the parser would have sufficient evidence to label an RC as “long” once the number of words in the RC has surpassed a certain threshold, such that a pre-RC boundary could be assigned even before the right edge of the RC is reached.

Finally, there are various points at which the parser could incorporate this prosodic information to make attachment decisions. For instance, the parser could make an initial attachment based solely on syntactic information, e.g., preferentially pursuing low attachment in accordance with the parsing principle of Late Closure, and then revise this initial attachment later on if necessary once prosodic structure has been assigned and consulted. Under this view, prosody would play the role of adjusting the final parse rather than directing first pass parsing. Alternatively, the parser could make an attachment decision based on both syntactic parsing principles and the first pass prosody. Finally, the parser could leave the attachment decision underspecified until later information - including prosodic structure - is available (Swets, Desmet, Clifton, & Ferreira, 2008).

To make this more concrete, consider once again the Visibility First Hypothesis, according to which the parser prioritizes attachment to whichever attachment site is most visible. The Visibility First Hypothesis will make different predictions about whether the parser should prefer high versus low attachment depending on whether it is basing the attachment decision on first pass prosody or the final default prosody. Consider again the sentences in (16b). In Chapter 3 I will walk through the principles that likely guide assignment of prosodic structure in more detail; for now, I stipulate which prosodies will be assigned at various points in the sentence. When the parser has only seen the first noun of the complex NP structure, as in (17), it will likely attempt to place this noun into a prosodic phrase with the preceding subject and verb. Upon encountering the PP *of the senior* in (18), the parser places this PP in a new prosodic phrase. Thus, when the parser encounters the relative pronoun *who*, *the senior* is more visible than *the friend*, as indicated by boldface, because the former is in the prosodic phrase currently being processed, while the latter is separated by a prosodic boundary. At this point, according to the Visibility Hypothesis, attachment of the RC to *the senior* should be preferred because this noun is more visible. Once the parser has finished the RC and determined it is short, as in (19), the second noun and the RC are still in a prosodic phrase together to the exclusion of the first noun. For this final prosody, *the senior* is still more visible than *the friend*, so again the Visibility First Hypothesis predicts that attachment to N2 will be favored. Thus, in the case of a short RC, the first pass prosody in (18) and the final default prosody in (19) would line up.

(17) The professor saw the friend...

Prosody: (The professor saw the friend)

(18) The professor saw the friend of the senior who...

Prosody: (The professor saw the friend) (of **the senior** who...)

(19) The professor saw the friend of the senior who studied.

Prosody: (The professor saw the friend) (of **the senior** who studied).

The situation is different when the RC is long, however. For the sentence with a long RC, the first two steps in (17) and (18) would be identical: at the start of the RC, the first pass prosody has *the senior* more visible than *the friend*, and the VFH predicts a preference for attachment to *the senior* at this point in time. However, once the parser has encountered the rest of the sentence, as in (20), it will have realized that the RC is long. The long RC encourages a boundary between *the* *senior* and the relative pronoun to achieve balance, and the boundary between the two nouns in the complex NP structure is removed. Thus, this is a case where the first pass prosody and the final default prosody conflict, because only the latter places the two nouns in the same prosodic phrase. Based on this final (default) prosody for a complex NP followed by a long RC, the Visibility Hypothesis would predict that the parser does not favor attachment to either noun, at least on the basis of the prosodic structure, because both nouns are equally visible. This will result in a greater rate of high attachment interpretations compared to the short RC in (19), because in the latter low attachment is preferred due to the greater visibility of the second noun.

(20) The professor saw the friend of the senior who studied for hours in the library

Prosody: (The professor saw the friend of the senior) (who studied for hours in the library.)

As mentioned earlier, these potential conflicts between first pass prosody and the final default prosody raise questions about when prosodic information is used and when attachment decisions are made. For now, I make the assumption that the prosodic information is always used and that the Visibility First Hypothesis governs the influence of prosodic information on both first pass and reanalysis decisions. Thus, in the case of the long RC, an attachment is made at the relative pronoun, and on the basis of the VFH and the prosody in (19), the parser is more likely to attach to *the senior*. However, once the parser reanalyzes the prosody as in (20), it will also readjust its attachment choices; since both nouns are now equally visible, the parser may sometimes shift away from its initial low attachment decision.

Although I have described this process as a serial parser that makes initial commitments and then revises them, it is also possible to frame this account in terms of a parallel parser that redistributes probability among different parses. In this case, at the relative pronoun in (19), the VFH predicts that the parser would allocate more probability to the low attachment parse than to the high attachment parse, because of the greater visibility of the second noun. However, the VFH predicts that the parser would reallocate some of this probability back to the high attachment parse once it has arrived at the final default prosody in (20), because now both nouns are equally visible.

To summarize, while previous work has successfully established that relative clause length is linked to implicit prosody and affects attachment height, the theory of incremental implicit prosodic assignment is still underspecified. This problem extends beyond relative clause attachment: for any construction, a theory of implicit prosody must ultimately specify (i) at which point the parser has sufficient information to assign a prosodic boundary in a given position, and (ii) when the existence of said boundary affects syntactic parsing and interpretation. Much previous work has been unable to address these questions about timing because they have primarily collected end-of-sentence interpretation data, which cannot shed light on incremental processing, and because they have mostly (but not exclusively) focused on manipulating the length of the relative clause, which necessarily occurs *after* the position of the hypothesized pre-RC boundary.

1.4 Goals of the Dissertation

With the preceding discussion, I hope to have demonstrated that the ample work documenting the influence of various prosodic features in silent reading belies the lack of theorizing about the incremental assignment of implicit prosody and the issues related to the potential divergence of first pass prosody and the final default prosody. I do not attempt to develop a fully specified theory of prosodic structure assignment in this dissertation. However, taking such a theory to be the eventual goal of work in this area, this dissertation lays the groundwork for addressing the following questions:

(21) Which methods are suitable for studying the online assignment of implicit prosodic structure?

(22) Where and how do implicit prosodic effects show up in online reading measures?

(23) What might an incremental prosodic parser look like? What principles would guide the parser's decisions, and how would they be deployed incrementally?

(24) How do features of the implicit prosodic representation affect syntactic parsing and interpretation? How do they interact with other properties of the stimulus, such as focus?

Chapter 2 of the dissertation is a methods chapter. Eye-tracking while reading would be the ideal method for studying the incremental assignment of implicit prosody, because this task most closely approximates natural reading situations and has measures that reflect both early and late processes. However, I was unable to run eye-tracking studies for this dissertation because of the COVID-19 pandemic, and had to rely on alternative methods. To that end, Chapter 2 addresses the viability of the Maze (Forster, Guerrera, & Elliot, 2009; Witzel, Witzel, & Forster, 2012), a reading task in which participants proceed through a sentence word-by-word, for the study of implicit prosody. At each position in the sentence, participants are presented with two possible continuations of the sentence, and they must pick which word forms the best continuation. This method has been argued to have several advantages over self-paced reading (SPR) (Boyce, Futrell, & Levy, 2020; Witzel et al., 2012), but it is also more artificial than self-paced reading and eye-tracking, and it is unclear whether the task demands, including pausing after each word to decide which of the two options is the best continuation, would interfere with the assignment of implicit prosody.

In the first part of Chapter 2, I show that the Maze is sensitive to metrical structure, replicating Breen and Clifton's (2011, 2013) experiments on homographs in which they find a greater processing cost when reanalysis requires both a change in lexical category and a change in the position of primary stress. I replicate this finding in both the Maze and SPR, while confirming previously reported advantages of the Maze over SPR, including more localized and larger effects.

In the second half of Chapter 2, I extend this methodological investigation to phrasal prosody. Previous research has suggested that the Maze introduces hyper-incrementality, a tendency to close clausal constituents wherever possible (Witzel et al., 2012). One way of explaining this hyper-incrementality is that stopping at each word in the Maze to determine the best continuation introduces a boundary-heavy prosody, and that these boundaries are then interpreted as reflecting the end of a syntactic constituent. If so, I would be unable to use the Maze to study phrasal prosody, because the task would encourage an implicit prosody that differed from the default prosody assigned in more typical reading situations. In Experiments 3 and 4, I replicate the NP/Z garden path in the Maze and SPR, showing that it cannot be the case that the Maze has introduced a boundary-heavy prosody, because such a prosody would have circumvented the garden path. In Experiments 5 and

6, I replicate the finding that increasing RC length leads to more high attachment responses in both the Maze and SPR Hemforth et al. (2015). Under the assumption that the RC length effect reflects implicit prosody, the replication of this effect shows that the Maze is sensitive to implicit prosodic effects that are found in more typical reading situations. Together, these findings support the use of the Maze to study phrasal prosody.

Chapter 3 presents a toy model of how a parser would incrementally assign prosodic structure to several well-known sentence structures. The discussion provides an overview of the grammatical constraints on the syntax-prosody interface and how these could be incorporated into an incremental parser. The chapter also highlights the frequent discrepancies between first pass and default prosody, motivating the need for further investigation of incremental structure building.

Chapter 4 investigates the incremental assignment of implicit prosodic boundaries. I present a series of experiments in the Maze in which I manipulate the length of the NP preceding the RC in order to test for online effects of implicit prosody in addition to collecting end-of-sentence interpretation data. Overall, these experiments provide suggestive evidence about where implicit boundaries might show up in response times, though I also lay out reasons to be skeptical of relying on response times to draw inferences about implicit prosody. However, by laying out the logic of the study, the discussion highlights the issues that will be important to consider in future work on the incremental assignment of prosody.

The experiments also complicate our understanding of length-based effects by showing that both changes in constituent length and the content of the additional material influence participants' interpretation of the sentence. I argue that modification in these studies plays a similar role to pitch accents, drawing focus to modified constituents. The results suggest that modification attracts attachment to a modified noun, but only when it is in the same implicit prosodic phrase as another potential attachment site. I argue that this provides support for the Visibility First Hypothesis (VFH), according to which prosodic visibility takes priority in determining attachment preferences, while modification only affects the parser's choices when two attachment sites are equally visible. I then outline several follow-up studies to test the Visibility First Hypothesis in future work.

Chapter 5 concludes.

Chapter 2

Implicit prosody in the Maze and Self-Paced Reading

2.1 Background

In Chapter 1, I identified the potential differences between first pass implicit prosody and the final "default" prosody as an area that deserves further attention. Under ideal circumstances, eye-tracking while reading would be the method used to study the incremental assignment of implicit prosody: this method has various measures for both first pass reading and later reanalysis and reprocessing, which would be useful for disentangling first pass prosody from later revisions. Moreover, the method provides a closer approximation of natural reading situations than other tasks like self-paced reading (SPR) and the Maze.

Because of the COVID-19 pandemic, I was unable to run eye-tracking while reading studies for this dissertation, having to rely instead on SPR and the Maze. These tasks are artificial compared to typical reading situations, because participants can only see one word at a time, and do not have the ability to return to earlier parts of the sentence. In some ways, this property makes these tasks more similar to a listening task, and they may still be appropriate for studying prosodic processing; however, it also complicates the task of comparing first pass prosody and default prosody. Moreover, as I describe in more detail below, these methods introduce task demands that could plausibly impact the assignment of implicit prosody. To address this potential issue, this chapter is focused on methods: I replicate various implicit prosodic effects in both the Maze and SPR, justifying the use of these tasks in this area and laying the groundwork for a series of Maze experiments in Chapter 4 that aim to address the issue of first pass versus default prosody.

2.2 Reading Tasks

2.2.1 The Maze Task

The Maze is a task in which participants are presented with two words at a time, one of which forms a grammatical continuation with the preceding material (the *target*) and one of which does not (the *foil*). Participants must select the target in order to advance through the sentence; selecting the foil causes the trial to terminate. This task stands in contrast to eye-tracking while reading, in which an entire sentence is presented at once and readers' eye movements are recorded as they progress through the sentence, and SPR, in which participants are presented with one word at a time and press a button to proceed from one word to the next.

The Maze is thought to encourage incremental processing, because participants must take into consideration all preceding material in order to decide which of the two words is the correct continuation. In support of this view, previous comparisons of the Maze and SPR have found that the Maze provides more localized effects than SPR: effects in the Maze are often found at the disambiguating region, while effects in SPR are typically distributed across multiple spillover regions and are therefore more difficult to interpret (Boyce et al., 2020; Forster et al., 2009; Witzel et al., 2012). This suggests that, when feasible, the Maze may be a more reliable alternative to SPR.

However, there are some concerns with the Maze task that have potentially contributed to the fact that is less commonly used than eye-tracking while reading and SPR. The first is a practical concern: developing foils that are ungrammatical continuations, or at least highly implausible, is often quite difficult, and creating foils for an entire experiment is a labor-intensive task. Recently, Boyce et al. (2020) introduced the A(uto)-Maze, which uses natural language processing technology to automatically generate high surprisal foils. The introduction of the A-Maze has significantly facilitated the creation of foils, addressing this potential barrier to running Maze experiments.

Another concern relates to the nature of the task itself. On a continuum from most natural to least natural reading task, eye-tracking while reading is the most natural, as it most closely approximates typical everyday reading situations, while the Maze is the most unnatural because of the requirement that participants read and decide between two words at each point in the sentence. One could reasonably ask whether the Maze is sensitive to the same processes that underlie sentence comprehension in more natural reading situations. Despite this concern, many classic sentence processing effects have now been replicated in the Maze, including the subject relative clause (SRC) advantage (Forster et al., 2009), lexical ambiguity resolution (Forster et al., 2009), relative clause and adverb attachment preferences (Boyce et al., 2020; Witzel et al., 2012), the NP/S coordination ambiguity (Boyce et al., 2020), and the ambiguity advantage effect (Sloggett, Van Handel, Sasaki, et al., 2020). Together, these replications show that the Maze is sensitive to many of the same sentence processing effects found in eye-tracking and SPR experiments, which suggests that the Maze is a useful and reliable experimental method for psycholinguists despite the additional idiosyncratic task demands it introduces. Indeed, in addition to these replications, the Maze has been fruitfully applied to novel studies on the processing of ellipsis (Kroll, 2020), discourse (Sasaki, 2021), focus (Hoeks, Toosarvandani, & Rysling, 2021), and relative clauses (Rich et al., 2022).

Despite the potential of the Maze as an alternative to other reading tasks, I know of no studies that have attempted to replicate effects of implicit prosody in the Maze. However, it is plausible that the Maze task could interfere with implicit prosody in unique ways that are not addressed by previous replications of syntactic processing effects. For instance, pausing at each choice point to decide between a target and a foil word could cause participants to adopt a staccato prosody containing more boundaries than would be typical under more natural reading conditions. Suggestive evidence for this hypothesized boundary-heavy prosody comes from Witzel et al. (2012), who propose that the Maze introduces "hyper-incrementality," a tendency to close clausal constituents whenever possible.

Witzel et al. (2012) appeal to hyper-incrementality on the basis of the NP/S coordination ambiguity, an example of which is provided in (25). In this structure, there is a temporary ambiguity in (25a) in which the string *the jeweler and the salesman* could be interpreted either as an NP coordination serving as the direct object of the verb *shot*, or *the jeweler* could be analyzed as the sole object, in which case *and* would be analyzed as introducing the coordination of a new clause with *the salesman* as the subject. Typically, readers will first pursue the NP analysis, resulting in a slowdown in reading times at the verb *reported*, which disambiguates to the S-coordination parse (Boyce et al., 2020; Frazier & Clifton, 1997; Witzel et al., 2012). This garden path does not arise in (25b), where a comma after *jeweler* immediately disambiguates to S-coordination.

(25) a. The robber shot the jeweler and the salesman reported the crime to the police.

b. The robber shot the jeweler, and the salesman reported the crime to the police.

While Witzel et al. (2012) found evidence of this garden path in SPR and eye-tracking while reading experiments, they failed to find any evidence of a garden path in the Grammatical-Maze or the Lexical-Maze tasks. This is where hyper-incrementality comes in: if participants close clauses as soon as possible, then they would always interpret *jeweler* as the end of the clause and pursue an S-coordination parse, even in the absence of a comma. As such, Witzel et al. (2012) caution that the Maze may not be "appropriate for investigating clause/constituent closure commitments during online sentence processing."

Although they do not mention prosody, and appeal to hyper-incrementality based on the fact that the task encourages participants to fully process material in order to determine which word forms the best continuation, a non-mutually exclusive possibility is that this this tendency is caused by the way the Maze affects implicit prosody. If implicit prosody in the Maze contains more boundaries than in other reading tasks, and these boundaries are construed as cues to syntactic structure, then the higher likelihood of a boundary after the word *jeweler* in this task would also encourage closing the clause early, leading to apparent hyper-incremental parsing. Thus, it is possible that their finding should also caution us against adopting the Maze not only for studies of closure commitments, but for studies of implicit prosody more generally.

There are some caveats, however. To truly show that the Maze encouraged hyper-incrementality, leading participants to favor the S parse, it would have been necessary to include conditions showing that participants also struggle when the sentence is disambiguated to the NP parse, since a hyper-incrementality account predicts that the NP parse should now be dispreferred. More recently, Boyce et al. (2020) have successfully replicated the NP/S coordination ambiguity garden path in the Maze, which also raises doubts about the finding reported by Witzel et al. (2012); with only 24 items and 32 participants, it is possible that the G-Maze and L-Maze studies failed to find an effect due to a lack of power.

Beyond the possibility of a staccato prosody, which I have connected to hyper-incrementality, there are other ways in which the Maze could influence the assignment of prosody. In the Maze, participants may repeat the sentence to themselves at various points to help decide between the target and the foil. In the process of this repetition, the first pass prosody could get smoothed into a more natural contour, one that is closer to the final "default" prosody. If this is the case, then concerns about a staccato prosody would be unfounded. However, this scenario could still result in an assignment of prosody that differs from what occurs in everyday reading: the smoothing over of the contour caused by repetition could cause the parser to arrive at the more natural default prosody earlier than it otherwise would. Yet another possibility is that repeating the preceding material could cause participants to reinforce the first pass prosody. This greater commitment to the first pass prosody could make it more costly to revise to the default prosody for the final sentence later on.

Thus, there are many plausible ways that the Maze may or may not interfere with implicit prosody. This uncertainty about whether and how the Maze would impact implicit prosody would complicate the interpretation of results from studies using this method: it would be difficult to conclude whether an effect of an implicit prosodic manipulation in the Maze were task-specific or informative about implicit prosody more generally. The primary purpose of this methods chapter is to show that such concerns are unfounded: the Maze is sensitive to various implicit prosodic effects, suggesting that it is an appropriate method to use as an alternative to eye-tracking.

2.2.2 Self-paced reading

In self-paced reading, participants read sentences one word at a time, pressing the space bar to move from one word in the sentence to the next. Once the space bar is pressed, participants cannot return to earlier parts of the sentence. Anecdotally, SPR often lends itself to many different strategies that could potentially affect depth of processing in general and the assignment of implicit prosody in particular. Some participants report tapping the space bar to proceed through the experiment as fast as possible, ostensibly to help them better remember the sentence for the comprehension question. In this case, the SPR task may encourage a more superficial scanning strategy, such that participants do not process the sentence as deeply, or process words at a delay; this would also explain why effects often show up in spillover regions. With regard to implicit prosody, superficial scanning may lead to a less detailed implicit prosodic representation than under normal reading conditions, and tapping through a sentence quickly may lead to fewer prosodic boundaries and/or different phrasings than would typically be found in silent reading. For this reason, understanding how and when implicit prosodic effects manifest in SPR is another important step to guide future research in this area.

Beyond implicit prosody, there are other concerns about using SPR instead of the Maze. A common finding across studies comparing these tasks is that the Maze may be a more reliable method than SPR: in SPR, effects often show up in a spillover region, if they show up at all (Boyce et al., 2020; Sloggett, Van Handel, Sasaki, et al., 2020; Witzel et al., 2012). In addition to questions of task sensitivity, it can be more difficult to interpret SPR data, because it is not always clear whether effects that show up *after* the critical manipulation were actually caused by this manipulation or by some other unidentified difference between conditions. Thus, it is worth comparing SPR and Maze on additional phenomena: showing that advantages of the Maze, such as more localized effects, are also found in studies on implicit prosody would strengthen the argument in favor of using the Maze despite the more artificial nature of this task.

To determine to what extent the Maze and SPR are viable methods for studying implicit prosody, this chapter presents a series of experiments replicating well-known implicit prosodic effects in both tasks. I address effects of metrical prosody in Experiments 1 and 2, and effects of phrasal prosody in Experiments 3 through 6. Although both tasks are shown to be sensitive to implicit prosody, I find many of the same advantages previously reported for the Maze, such as more localized and larger effects. The results of this methods chapter support the use of the Maze in Chapter 4 and add to a growing body of work comparing different reading tasks (Boyce et al., 2020; Forster et al., 2009; Sloggett, Van Handel, Sasaki, et al., 2020; Witzel et al., 2012).

2.3 Implicit metrical prosody

Breen and Clifton (2011) showed evidence that metrical structure is assigned during silent reading. In an eye-tracking study, participants read limericks designed to generate expectations about the location of strong and weak syllables. They then manipulated whether the stressed syllable of a target word (e.g., *preSENT*) appeared in an expected strong syllable position (26a) or a weak syllable position (26b), as determined by the meter of the limerick. If participants generate a metrical structure during silent reading, Breen and Clifton reasoned that the presence of an unstressed syllable in an expected strong position, as in (26b), would require revision to the metrical structure, which would cause processing difficulty. Indeed, Breen and Clifton found that participants were less likely to skip the critical word and had longer reading times in (26b) relative to (26a). A later ERP study showed that the metric violation induced by reading these limericks generated ERPs consistent with those found in studies of metric violations in explicit prosody, providing further support for the claim that readers were generating a metrical structure during silent reading (Breen et al., 2019).

(26)

	(W	S)	(W	W	S)	(W	W	S)	(W)
a.	There	once	was	a	cle	ver	young	gent	
	who	had	a	nice	talk	to	**pre**	**SENT**	
b.	There	once	was	a	penn	i	less	pea	sant
	who	went	to	his	mas	ter	to	**pre**	**SENT**

In a second experiment, Breen and Clifton (2011) showed that readers assign metrical structure even outside of the highly constraining context provided by limericks. They used homographs to create temporarily ambiguous sentences: the string *the brilliant report* could be an adjective + noun (i.e., a report of high quality), as in (27), or a noun + verb (i.e., brilliant people are reporting), as in (28). All of the adjective-noun homographs were chosen to be biased towards the adjective reading, to ensure that readers initially analyzed the noun-verb homograph as a noun, requiring syntactic reanalysis from the structure A+N to the structure N+V in (28) when the following string *the best ideas* indicates that the N+V analysis is required. As expected, they found a processing cost reflecting syntactic reanalysis when the sentence disambiguated toward the N+V reading.

(27) The [brilliant]$_{\mathrm{A}}$ [report]$_{\mathrm{N}}$ was accepted at the prestigious conference.

(28) The [brilliant]$_{\mathrm{N}}$ [report]$_{\mathrm{V}}$ the best ideas from the things they read.

Breen and Clifton also varied whether the noun-verb homographs had different stress patterns (e.g., noun *ABstract* vs. verb *abSTRACT*), as in (29) and (30). They reasoned that if readers assigned implicit metrical structure in silent reading, then the switch from A+N to N+V would incur a greater processing cost with the homograph *abstract* in (30) than with the homograph *report* in (28), as only (30) would require both syntactic and metrical reanalysis. As expected, they found greater processing difficulty when both types of reanalysis were required, which was reflected in longer first fixation and first pass times on the ambiguous homograph in this condition. This additional cost for metrical reanalysis supports the claim that readers construct implicit prosodic representations in silent reading, and that these representations include lexical stress.

(29) The [brilliant]$_{\mathrm{A}}$ [abstract]$_{\mathrm{N}}$ was accepted at the prestigious conference.

(30) The [brilliant]$_{\mathrm{N}}$ [abstract]$_{\mathrm{V}}$ the best ideas from the things they read.

Breen and Clifton (2011) found that the extra cost of metrical reanalysis appeared on the ambiguous word itself, prior to when readers fixated the disambiguating region. They reasoned that parafoveal preview allowed readers to identify the next word in the disambiguating region, which provided sufficient information to disambiguate the part of speech of the ambiguous noun-verb homograph and initiate reanalysis while they were still fixated on this word, explaining why the effect showed up before the disambiguating region was fixated. In a subsequent study, Breen and Clifton (2013) tested this account by using a boundary change paradigm that prevented parafoveal preview of the disambiguating region. In this follow-up experiment, the extra cost of metrical reanalysis appeared after the disambiguating region, rather than on the ambiguous homograph. This finding supported the claim that the longer first fixation and first pass times on the ambiguous word in the original experiment were not spurious but rather reflected metrical reanalysis triggered by information from parafoveal preview.

The present study attempts to replicate this finding in the Maze and SPR in order to establish that these tasks are sensitive to implicit prosodic effects. Breen & Clifton's homographs experiment was chosen because it shows a clear and replicable effect of metrical structure in eye-tracking while reading. Thus, one can reasonably expect that this effect should show up in alternative reading tasks, if these tasks are also sensitive to implicit metrical structure. Moreover, word-by-word paradigms like the Maze and SPR could provide additional support for the parafoveal preview account of the timing of the effect in the original experiment: because preview of the next word is not possible in these tasks, the effect of metrical reanalysis should show up on or after the disambiguating region, as in Breen and Clifton's follow-up study. The Maze vs. SPR comparison also allows us to verify that the purported advantages of the Maze over SPR, such as more localized effects, extend to experiments on implicit prosody, which would further motivate the adoption of the Maze over SPR in spite of the more artificial task.

2.4 Experiment 1: Homographs in the Maze

2.4.1 Methods

2.4.1.1 Materials

The materials consisted of the 32 items from Breen and Clifton (2013) and eight additional items created for the replication study in order to increase power; an example item is provided in (31). Each sentence began with an ambiguous string consisting of a determiner, an adjective-noun homograph, and a noun-verb homograph, followed by a disambiguating region. In the Noun condition, the disambiguating region usually began with a verb; in the Verb condition, this region usually began with a determiner. Participants saw only one condition for each item (cf. Breen and Clifton's (2013) experiment, in which participants saw one Noun condition and one Verb condition for each item). A full list of experimental items is provided in Appendix B. In addition to the 40 experimental items, there were 80 filler items: 48 items came from an experiment on the NP/Z garden path, described in Section 2.7, and the remaining 32 were from an experiment on retrieval interference. All sentences were followed by a yes/no comprehension question.

(31) Example item

Part of Speech	Stress Pattern	Sentence
Noun	Alternating	The secretive **project** was closely guarded by the military.
	Non-alternating	The secretive **design** was closely guarded by the military.
Verb	Alternating	The secretive **project** an image of mystery and privacy.
	Non-alternating	The secretive **design** an image of mystery and privacy.

Based on the findings of Breen and Clifton (2011, 2013), I predicted a main effect of Part of Speech (POS) on the disambiguating word (*was/is*) immediately following the noun-verb homograph (*project/design*), such that response times are longer in the Verb condition, reflecting the cost of syntactic reanalysis. I also predicted an interaction of POS and Stress Pattern, such that the Verb cost would be greater in the Alternating condition compared to the Non-alternating condition, because in this condition the syntactic analysis requires a concomitant metrical reanalysis. I also predicted that the main effect of POS and the interaction of POS and Stress Pattern would show up in the Maze error rates at the disambiguating region, such that participants would incorrectly choose the foil instead of the target at a greater rate in the Verb and Non-Alternating conditions, under the assumption that the additional processing difficulty of the garden path would interfere with participants' ability to arrive at the correct parse and select the right continuation.

The foils were generated using Boyce et al.'s (2020) A-Maze, which automatically generates high surprisal (and therefore unlikely) continuations at each point in the sentence. The foils produced by the A-Maze were checked by hand and revised in cases in which the foil could have been a grammatical continuation and when the same foil was used frequently across multiple items. Because this is a methods chapter, the present study also introduced a between-subjects manipulation that varied whether the stress pattern of the foils matched the stress pattern of the targets, as an exploratory methodological question. To create the mismatched foils, I took the foils generated by the A-Maze, identified all cases where the target and the foil shared the same stress pattern, and manually generated a new mismatched foil. Similarly, to create the matched foils, all cases where the target and the foil had different stress patterns were identified, and a matched foil was manually generated to replace the original foil. After running the version of the experiment with matched foils, it was found that a small percentage of the foils were actually mismatched; these foils were changed, and a third version of the experiment with perfectly matched foils was run. Thus, there were three subexperiments that differed in foil type: Stress Mismatched, Mostly Stress Matched, and Stress Matched.

The choice to compare different foil types was motivated in part by recent work investigating to what extent foil type affects the sensitivity of the Maze task; for instance, Boyce et al. (2020) found that the G(rammaticality)-Maze, in which the foil is a real word but an ungrammatical continuation, and all foils are generated by hand, was more powerful than the L(exicality)-Maze, in which the foil is a nonce word, and that the G-Maze and A-Maze produced comparable results. Similarly, Sloggett, Van Handel, and Rysling (2020) compared manually generated ("handmaze," comparable to G-Maze) and A-Maze-generated foils and found that the foil source did not impact the effect of interest, although participants overall made more errors in the A-Maze, particularly in earlier sentence positions. Though the reasons for differences across various Maze tasks are not of great theoretical interest, the outcomes of these studies are helpful in determining best experimental practices in the Maze, which is still relatively new. For instance, although creating nonce-words in the L-Maze is easier than generating ungrammatical continuations in the G-Maze, the greater power of the G-Maze is likely worth the extra effort. In contrast, it is much easier to generate foils in the A-Maze compared to the G-Maze, and the comparable results in the A-Maze and G-Maze/handmaze in previous studies suggests that this more efficient foil generation is not marred by a cost in experimental sensitivity.

In the context of the present study, there are various ways in which the stress pattern of the foil could plausibly impact participants' responses in the Maze. For instance, if participants consistently see targets and foils with matching stress patterns, it may be easier for readers to assign a metrical structure, because they do not have to shift between different stress patterns when testing out the relative goodness of fit of the target and the foil; in this case, participants may have a more stable metrical structure, making them particularly sensitive to the unexpected stress pattern of the homographs in the Verb, Alternating condition and resulting in a stronger effect. In contrast, consistently seeing targets and foils with divergent stress patterns may make it harder for readers to keep track of the overall metrical structure and/or make readers less likely to pay attention to metrical structure because they have to continuously shift between conflicting stress patterns when deciding which word serves as the best continuation. If mismatching targets and foils cause participants to pay less attention to metrical structure, this could result in a weaker or nonexistent cost of metrical reanalysis. Foils with mismatched stress patterns could also facilitate metrical reanalysis of the stress-alternating noun-verb homographs, if participants get used to having to shift between different metrical structures.

The stress pattern of the foil could also affect the results via priming. For instance, if the stress pattern of the foil matches the stress pattern of the target, then reading the foil before the ambiguous target could prime the target stress pattern; even if the foil is read after the ambiguous target, its stress pattern could make the target stress pattern more readily available. If such a priming effect existed, it could reduce or eliminate the cost of metrical reanalysis when the stress pattern of the foil matches the target by helping the reader arrive at the required pattern. In contrast, a mismatched foil could prime the wrong stress pattern, leading to a greater cost of metrical reanalysis. Support for the possibility that the foil could prime implicit prosody comes from Wilkenfeld (1985), described in Breen (2014), who showed that speakers could be primed to produce noun-verb homographs with the dispreferred weak-strong metrical pattern after silently reading a list of words that shared this pattern, and from Ashby and Rayner (2004), who showed that syllabic information in parafoveal preview can facilitate lexical access.

To reiterate, the foil type comparison is not driven by a theoretical question, and I am not committed to a particular hypothesis about how foil type should affect performance in this experiment; there are multiple logically possible ways in which the stress pattern of the foil could, but does not have to, affect performance. Following previous work comparing different types of foils, my intention is simply to test whether and how the type of foil affects the sensitivity of the Maze task. If the stress pattern of the foil interacts with the effect of interest in the present study, then researchers would need to be concerned about the stress pattern of foils when running implicit prosody experiments in the Maze in the future, placing yet another constraint on the generation of foils. If instead the type of foil does not interact with the effect of interest, as in Sloggett, Van Handel, and Rysling's (2020) comparison of the A-Maze and the handmaze, then there is no need to control for the stress pattern of foils in future studies. Moreover, by showing that foil properties do not affect the assignment of implicit prosody, I can strengthen the argument in favor of using the Maze.

2.4.1.2 Participants

196 native English speakers participated in the experiment. 70 subjects were recruited from Prolific Academic and were paid $8 for participating. These subjects were run on the version of the experiment with the Mostly Stress Matched foils. Of these 70 subjects, ten were excluded for poor performance, which was defined as failing to complete at least 50% of the sentences in the Maze. The other 126 subjects were undergraduates at UC Santa Cruz and received course credit for their participation. 63 of these participants were run on the version of the experiment with Stress Matched foils; of this group, two participants were excluded for poor performance, and one was excluded because they participated after the target of 60 subjects had already been reached. The remaining 63 participants were run on the Stress Mismatched version of the experiment; of this group, three participants were excluded for poor performance.

2.4.1.3 Procedure

The experiment was conducted online and administered on Ibex Farm. Each trial began with the first word of the trial appearing on the left side of the screen, above the letter “e” and a series of x’s appearing on the right side of the screen, above the letter “i” as shown in Figure 2.1. To start the trial and choose the first word, participants pressed the “e” key on the keyboard. Each subsequent word in the sentence was presented as a choice between the correct target word and the incorrect foil, as shown in Figure 2.2. The position of the target and the foil was randomized. Participants pressed the “e” key to choose the word on the left and the “i” key to choose the word on the right. The trial terminated if participants incorrectly chose the foil. Participants were given an opportunity to take a short break after every 20 trials.

Figure 2.1: Example of the presentation of the first word of a trial

Figure 2.2: Example of the presentation of a foil and target

At the end of each trial, participants were presented with a yes/no question testing their comprehension of the preceding sentence, as shown in Figure 2.3. Participants pressed the “e” key to respond “yes” and the “i” key to respond “no.” The comprehension question was still presented if participants made an error in the Maze task that caused the trial to terminate before the end of the sentence. Only data from the trials on which participants successfully completed the Maze for the entire sentence was included in the analysis of comprehension question accuracy.

Figure 2.3: Example of the presentation of a comprehension question

2.4.2 Results

2.4.2.1 Response times

Only response times for those cases in which the participant correctly chose the target word, and not the foil, were included in the analysis. Next, extreme values were trimmed by eliminating all response times that were less than 100ms or that exceeded 5000ms, which resulted in a loss of 1.4% of observations at the critical disambiguating word (*the*/*was*). The response times for the three word ambiguous string at the beginning of the sentence, the disambiguating word, and the following two spillover words are shown in Figure 2.4. This plot shows longer response times at the disambiguating word in the Verb condition, with particularly long RTs in the Verb, Alternating condition. The mean RTs and standard error at this word are reported in Table 2.1.

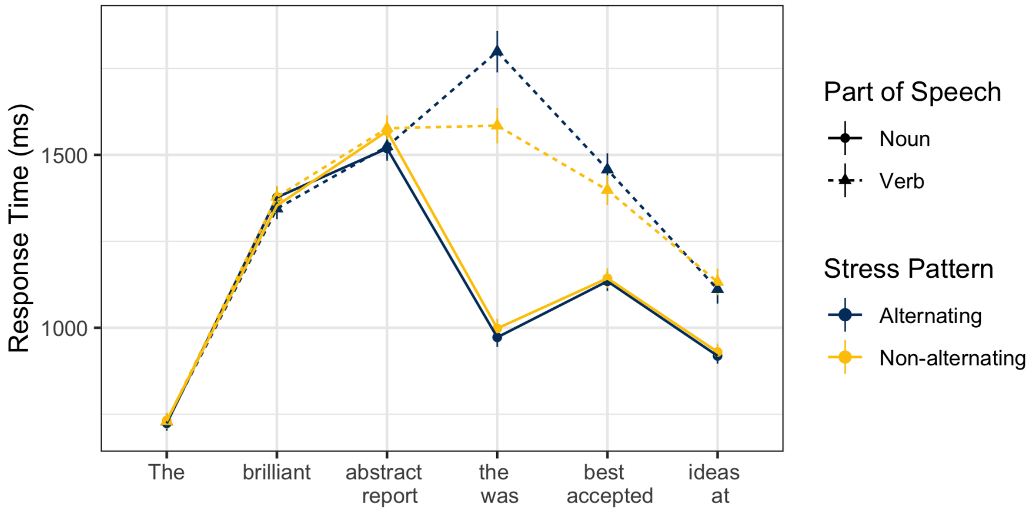


Figure 2.4: Mean response times by word in Experiment 1. Error bars show a 95% confidence interval.

POS	Stress	RT (ms)
Noun	Alternating	972 (14)
	Non-alternating	998 (14)
Verb	Alternating	1799 (31)
	Non-alternating	1585 (26)

Table 2.1: Mean response times (ms) at point of disambiguation in Experiment 1

A Bayesian mixed effects model was fit to the response times at the disambiguating word (*the*/*was*) with POS, STRESS, and their interaction as fixed effects and the maximal random effects structure. There were main effects of POS (765, [614, 922]), with longer RTs in the Verb condition, STRESS (97, [44, 150]), with longer RTs for Alternating homographs, and their interaction (250, [94, 408]), with the longest response times in the Verb, Alternating condition.

2.4.2.2 Maze error rates

The rate of errors made in the Maze task were analyzed next. Figure 2.5 shows how many errors were made at each word in the first half of the sentence as a percentage of the total errors made across the 40 experimental items. Impressionistically, the error rates mirror the response times, with the largest number of errors made in the Verb, Alternating condition.

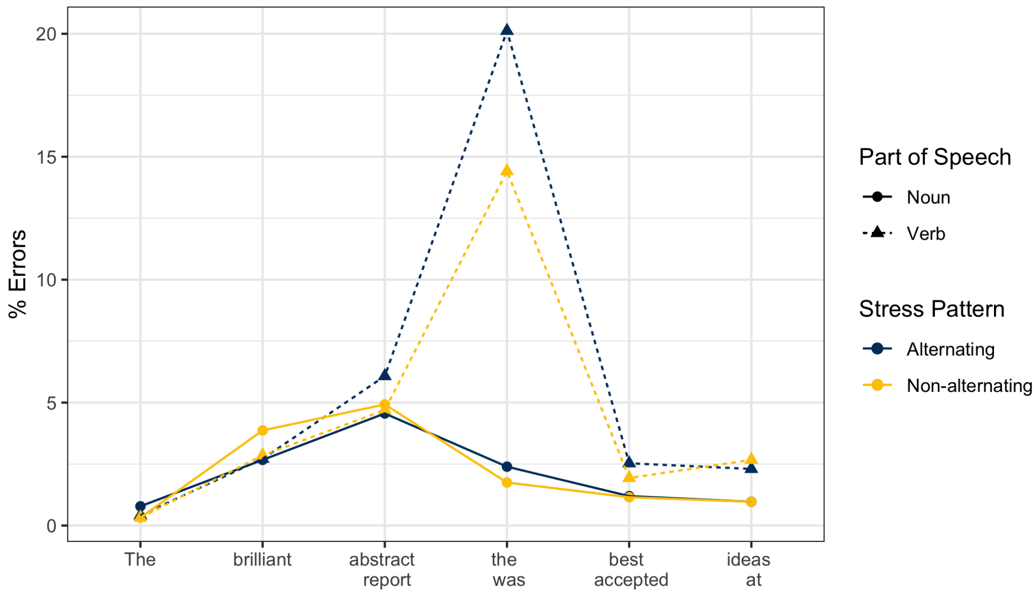


Figure 2.5: % of Maze errors by position in Experiment 1

Figure 2.6 shows the mean error rate in each condition, considering all errors made at or after the point of disambiguation as a percentage of the total number of trials in which participants made it to the point of disambiguation in each condition. This measure allows us to see the relative difficulty of completing the Maze in each condition, without including those errors that were made prior to the experimental manipulation. A Bayesian logistic mixed effects model was fit to the error rates with POS, STRESS, and their interaction as fixed effects and the maximal random effects structure. There was a main effect of POS (1.93, [1.61, 2.26]), with more errors in the Verb condition, and a main effect of STRESS (.26, [.07, .46]), with more errors for Alternating homographs. I did not find a significant interaction of POS and STRESS (.30, [-.1, .69]).

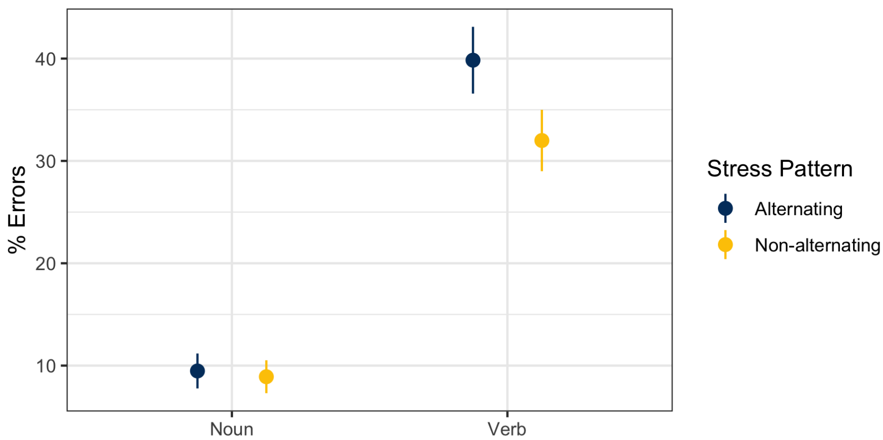


Figure 2.6: Mean error rates at or after the point of disambiguation in Experiment 1. Error bars show a 95% confidence interval.

Figure 2.7 shows the mean error rates at the disambiguating word as a percentage of the total number of observations at this position for each condition. This measure allows us to pinpoint the relative difficulty of each condition at the point of disambiguation in particular. A Bayesian logistic mixed effects model was fit to the error rates with POS, STRESS, and their interaction as fixed effects and the maximal random effects structure. I found main effects of POS (2.63, [2.07, 3.23]), with more errors in the Verb condition, and STRESS (.59, [.22, .99]). There was no significant interaction of POS and STRESS (.09, [-.68, .78]).

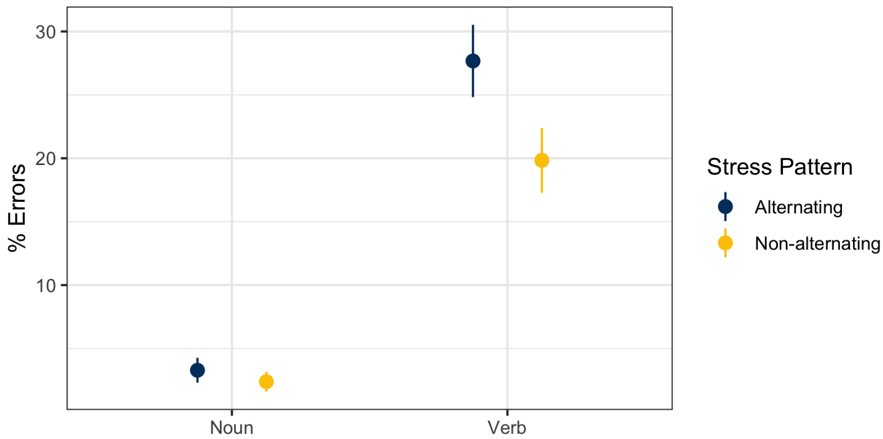


Figure 2.7: Mean error rates at the point of disambiguation in Experiment 1. Error bars show a 95% confidence interval.

2.4.2.3 Comprehension question accuracy

All trials on which participants made an error on the Maze task or did not respond to the question within the 60 second time limit were excluded from the analysis of the performance on the comprehension questions. This resulted in a loss of 2175 observations, or about 30% of the data, due to the high number of errors in the Maze task, particularly in the Verb condition. The overall accuracy across conditions is summarized in Table 2.2, with standard error computed over participant means. Performance on the comprehension questions was generally high; this is unsurprising, since the data come from those trials on which participants successfully completed the Maze.

POS	Stress	% Correct
Noun	Alternating	93 (.7)
	Non-alternating	92 (.8)
Verb	Alternating	84 (1.5)
	Non-alternating	85 (1.2)

Table 2.2: % correct responses to comprehension questions in Experiment 1

A Bayesian logistic mixed effects model of the probability of a correct answer was fit to the comprehension data with the same effects structure as the other models. Although the percentages suggest worse performance in the Verb condition, I failed to find a main effect of POS (-.96, [-1.91, 0]). I also failed to find a main effect of Stress (-.2, [-.72, .26]) and an interaction of POS and Stress (-.04, [-1.04, .93]).

2.4.2.4 Foil Type

To understand any potential role of foil type, response times across all items, including fillers, were analyzed. The average RTs at each position in the sentence across the entire experiment, separated by foil type, are shown with 95% confidence intervals in Figure 2.8. In general, it appears that response times were longer in the Mismatched sub-experiment, when the targets and foils had different stress patterns. One possibility, noted above, is that stress-matched foils allow for priming of the stress pattern from target to foil (and/or from foil to target), facilitating lexical access at each position in the sentence and resulting in overall faster response times. Another possibility is that for whatever reason the stress-matched foils were worse continuations than the stress-mismatched foils, which made it easier to reject them, leading to shorter response times. While the relative surprisal of the matched and mismatched foils was not controlled for, corresponding matched and mismatched foils always had the same part of speech and were both judged to be poor continuations of the sentence. Although there may be differences between the Mostly Stress Matched and Stress Matched conditions at some positions, it should be noted that these conditions only differed on a small percentage of foils, and that the Mostly Stress Matched sub-experiment was conducted with participants from Prolific, while the Stress Matched sub-experiment was conducted online with undergraduates from UC Santa Cruz. Thus, it is likely that any differences here are due to different populations rather than the foils themselves. Crucially, the difference between foil types appears to be one of magnitude, rather than a difference in the overall pattern of RTs.

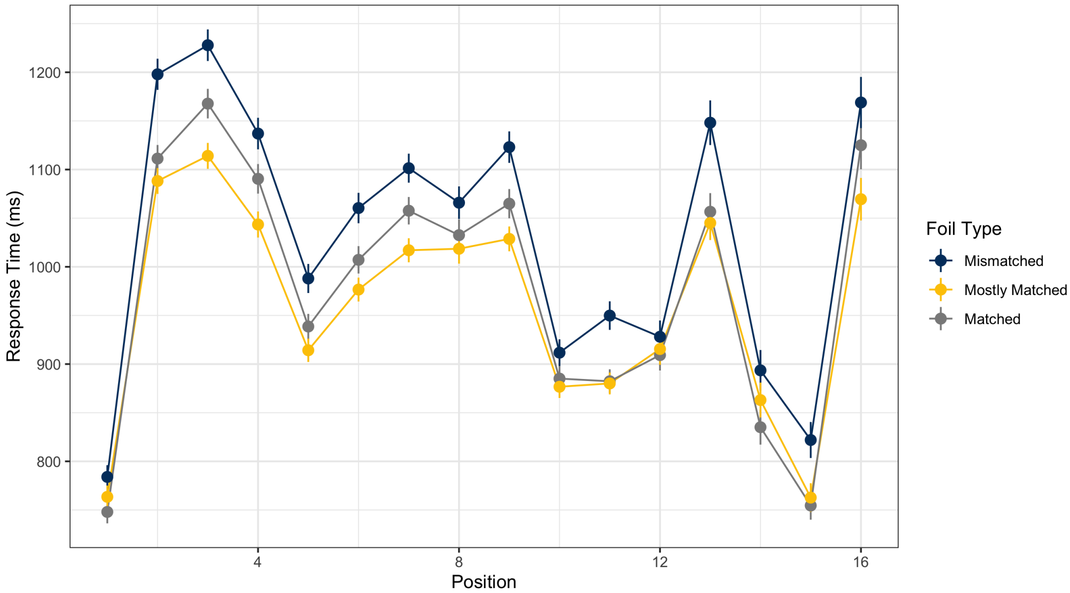


Figure 2.8: Mean response times at each position across all items, by foil type. Error bars show a 95% confidence interval.

Since the primary concern was whether the stress pattern of the foils would affect performance at the critical region in the homographs experiment, I next separated the crucial homographs region out by foil type, as shown in Figure 2.9. Impressionistically, there is the same qualitative pattern regardless of foil type, with the RTs at the point of disambiguation being longest in the Verb, Alternating condition. Separate models fit for each sub-experiment confirmed that the interaction of POS and STRESS was significant regardless of foil type.

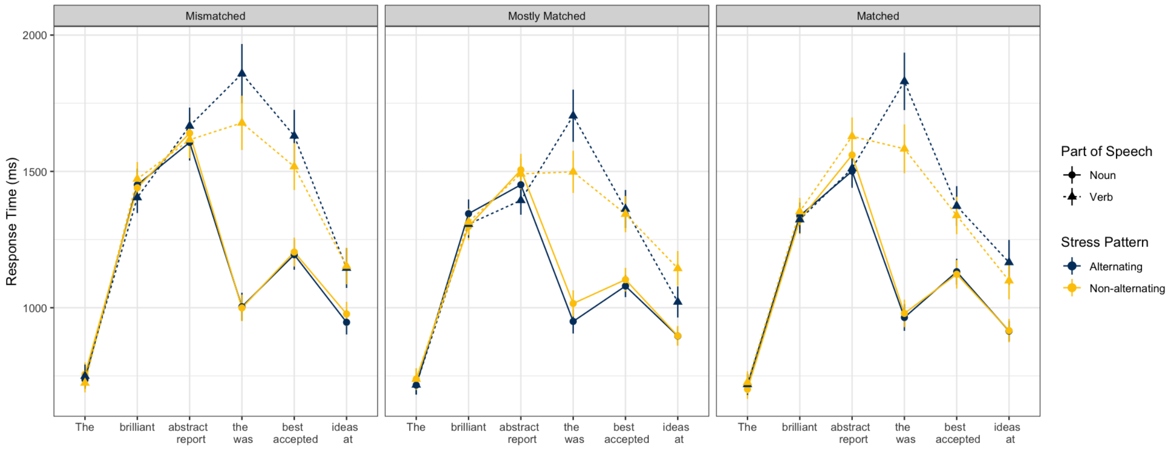


Figure 2.9: Mean response times at each position in the experimental items, by foil type. Error bars show a 95% confidence interval.

Again, I did not have a specific hypothesis about how foil type would interact with the experimental manipulation; rather, because this is a methods chapter, I wanted to check whether this potential source of variation would matter. The important point is that the stress pattern of foils does not appear to need to be controlled for in future studies, because it did not change the overall pattern of results in this experiment.

2.4.3 Discussion

This experiment successfully replicated the main findings from Breen and Clifton (2011, 2013): participants showed greater processing difficulty when the syntactic reanalysis of the noun-verb homograph from the noun to the verb reading was accompanied by a metrical reanalysis from a strong-weak to a weak-strong pattern.

The interaction of POS and STRESS was found in response times at the disambiguating word; this timing differs from what Breen & Clifton found, and I return to this point in the discussion of Experiment 2. Crucially, the interaction supports the claim that the Maze task is sensitive to implicit metrical structure, despite the artificial nature of the task. Thus, the Maze is a viable method for studying implicit prosody, at least for investigations of lexical stress, and this holds true regardless of the prosodic properties of the foils. Since the Maze is less resource-intensive than eye-tracking while reading, but potentially more sensitive and with more localized effects than self-paced reading, this is a promising result.

Moreover, the pattern of greater difficulty at the point of disambiguation in the Verb, Alternating condition appears to hold true, at least qualitatively, regardless of whether the foils matched or mismatched the stress pattern of the target words. This conclusion is similar in spirit to Sloggett, Van Handel, and Rysling's (2020) finding that both hand-generated and A-Maze foils produce qualitatively similar results. Because foil type did not impact the effect of interest in their study, they encouraged the use of the automatically generated A-Maze foils, which reduces a potential burden of developing materials for the Maze. Similarly, since this experiment provided evidence that the processing cost of metrical reanalysis was not affected by the stress pattern of the foils, I suggest that researchers do not need to control for the stress pattern of foils when conducting Maze studies on implicit prosody, which, based on my experience developing materials for this experiment, would be an onerous task and a potential barrier to the widespread use of the Maze in this area.

Finally, it is worth addressing the error rates: both in overall error rates and at the point of disambiguation, I found main effects of POS and STRESS but no interaction. It is unclear why there should be a main effect of STRESS in the absence of an interaction: by hypothesis, the difficulty at this position is due to reanalysis in the Verb condition, so I did not expect a difference between the Noun, Alternating and Noun, Non-alternating conditions. However, since the ambiguous homograph differed between the Alternating and Non-alternating conditions, it is possible that the target at the point of disambiguation was, for whatever reason, a worse fit after the Alternating homograph compared to the Non-alternating homograph, leading to more difficulty choosing the target over the foil. It is worth noting that Breen and Clifton (2013) also found a main effect of STRESS in the absence of an interaction in several measures: a significant effect on first fixation times, second-pass times, and probability of regression into the ambiguous homograph, as well as marginal effects on first-pass and go-past times on this word, which they suggest could be due to the non-alternating homographs being more frequent. They also found a marginal effect of STRESS on go-past times in the disambiguating region, but no interaction. Thus, a main effect of STRESS without an interaction is not unprecedented, and it is possible that lexical properties of the different homographs such as frequency also played a role in the main effect of STRESS on error rates, in particular if these properties affected the goodness of fit of the target after the homograph. Another possibility is that the main effect of STRESS on error rates at this position is spurious. Crucially, this effect does not undermine the interaction of POS and STRESS that was found in response times at the disambiguation point, which indicates greater difficulty when both syntactic and metrical reanalysis are required.

2.5 Experiment 2: Homographs in Self-Paced Reading

Having established that the Maze is sensitive to metrical prosody, I next report the results of a replication of the same Breen and Clifton (2013) experiment in the self-paced reading paradigm. This comparison adds to a growing body of work comparing the sensitivity of these two paradigms.

2.5.1 Methods

2.5.1.1 Materials

The same materials from Experiment 1 were used, with the exception that there was no foil manipulation, because the SPR task does not involve the presentation of foils.

2.5.1.2 Participants

65 undergraduate native English speakers at UC Santa Cruz participated in the experiment for course credit. Five participants were excluded because they participated after the target of 15 subjects for each of the four lists had been reached, leaving data from a total of 60 participants in the analysis.

2.5.1.3 Procedure

Each trial began with two dashes in the center of the screen. Participants pressed the space bar to make the dashes disappear and the first word appear. Sentences were presented one word at a time, and participants pressed the space bar to have the current word be replaced by the next word. The time spent on each screen before the participants pressed the space bar was recorded. Each sentence was followed by a "yes"/"no" comprehension question; participants pressed "e" to select "yes" and "i" to select "no."

2.5.2 Results

2.5.2.1 Response times

Extreme response times, defined as observations less than 100ms or greater than 5000ms as in Experiment 1, were eliminated from the data prior to analysis, resulting in a loss of .2% of observations from the entire homographs sentences, including .2% of observations from the critical disambiguating word and the first spillover word. The response times for the three word ambiguous string at the beginning of the sentence, the disambiguating word, and the following two spillover words are shown in Figure 2.10. As in the previous experiment, there is an apparent difference between the Noun and Verb conditions at the point of disambiguation. However, in contrast to the findings in the Maze, the longer response time in the Verb, Alternating condition compared to the Verb, Non-alternating condition does not appear to emerge until the first spillover word. The mean RTs and standard error at the disambiguating word and the first spillover word (*best/accepted*) are reported in Table 2.3.

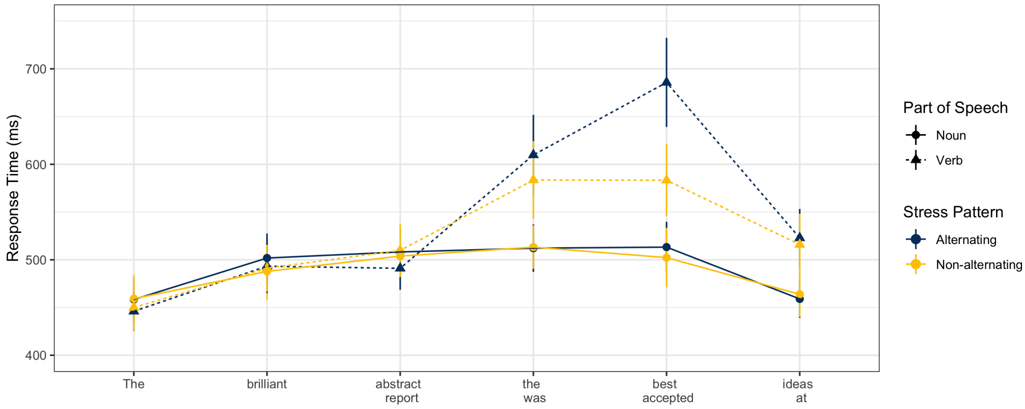


Figure 2.10: Mean response times by word in Experiment 2. Error bars show a 95% confidence interval.

POS	Stress	Disambiguation	Spillover
Noun	Alternating	512 (13)	513 (14)
	Non-alternating	513 (12)	502 (16)
Verb	Alternating	598 (21)	686 (24)
	Non-alternating	584 (21)	583 (19)

Table 2.3: Mean response times (ms) at point of disambiguation and spillover in Experiment 2

Bayesian mixed effects models were fit to the response times at the disambiguating word (*the*/*was*) and at the spillover word (*best*/*accepted*), with POS, STRESS, and their interaction as fixed effects and the maximal random effects structure. At the point of disambiguation, there was a main effect of POS (83, [28, 139]), indicating longer RTs in the Verb condition. I did not find a significant effect of STRESS

(13, [-19, 45]) or a significant interaction (28, [-38, 94]). At the first spillover word, there were main effects of POS (125, [68, 181]), with longer RTs in the Verb condition, STRESS (56, [11, 102]), with longer RTs for Alternating homographs, and their interaction (90, [4, 176]), with the longest response times in the Verb, Alternating condition.

2.5.2.2 Comprehension question accuracy

Participants responded to the comprehension question within the time limit on all trials, so all 2400 observations were included in the comprehension question analysis. The overall accuracy across conditions is summarized in Table 2.4, with standard error computed over participant means. As in Experiment 1, performance on the comprehension questions was high overall.

POS	Stress	% Correct
Noun	Alternating	93.5 (1)
	Non-alternating	92.2 (1.2)
Verb	Alternating	88.7 (1.6)
	Non-alternating	87.3 (1.4)

Table 2.4: % correct responses to comprehension questions in Experiment 2

A Bayesian logistic mixed effects model of the probability of a correct answer was fit to the comprehension data with the same effects structure as in previous models. There was a main effect of POS (-1.38, [-2.37, 0]), such that participants were less likely to answer correctly in the Verb condition. There was no effect of STRESS (-.21, [-.87, .38]), nor was there a significant interaction (-.08, [-1.46, 1.36]).

2.5.3 Discussion

As in Experiment 1, I found evidence that metrical reanalysis causes processing difficulty: there was an interaction of POS and STRESS on the spillover word, such that response times were longest when a stress-alternating homograph was disambiguated to a verb. This suggests that SPR is sensitive to lexical stress, despite the concern that a superficial scanning strategy encouraged by the task could lead to a less rich implicit prosody. On the basis of Experiments 1 and 2, I conclude that both the Maze and SPR are appropriate to study implicit metrical prosody.

An interesting methodological point is that the crucial interaction between POS and STRESS occurred at different points in each study: immediately at the point of disambiguation in the Maze, and at the first spillover word in SPR. These results stand in contrast to Breen and Clifton (2011), who found this interaction in first fixation and first pass reading times on the ambiguous homograph. This is clearly due to differences between methods. Breen and Clifton (2011) was an eye-tracking while reading study. As discussed above, they reasoned that their effect showed up on the ambiguous word because participants were able to extract information from parafoveal preview of the disambiguating region, meaning that they were able to begin metrical reanalysis while still fixated on the ambiguous homograph. Such preview is not possible in the Maze or SPR, where participants only see one word (or one target and one foil) at a time, so it is not surprising that the effect did not appear on the ambiguous word in Experiments 1 and 2. It is worth noting that the absence of the effect in this position is consistent with Breen and Clifton's appeal to parafoveal preview to explain their results.

The timing of the interaction also differed compared to Breen and Clifton (2013), who found the crucial interaction in go-past times in the final spillover region, after the point of disambiguation[1]. As described earlier, this was an eye-tracking while reading study that used a boundary change paradigm such that participants could not extract information about the homograph from parafoveal preview; this explains why the interaction no longer turned up on the homograph in the 2013 study. However, this still leaves the question of why the interaction occurred earlier in the Maze compared to the 2013 study. A likely reason is that the Maze forces incrementality due to the nature of the task, which requires participants at each choice point to take into consideration all material they have seen so far in order to decide which

[1]They also found marginal interactions in second pass times on the ambiguous homograph and on the disambiguating region.

word forms the best continuation. This leads participants to engage in syntactic and metrical reanalysis at the point of disambiguation, explaining why the response time slowdown appears immediately. In contrast, eye-tracking while reading allows readers to continue to read a sentence even before they have fully processed all preceding material; in this case, the task allows for a situation in which the full cost of syntactic and metrical reanalysis does not appear until further downstream.

While differences in the reading task can explain the differences in timing across studies, they also raise an important point about different tasks. While this study has added to the growing evidence that the Maze is sensitive to many of the same processes found in eye-tracking and SPR, it is also true that the artificial nature of the task means that results from the Maze do not necessarily reflect how these processes would unfold in more natural reading situations. That is, the Maze is an excellent tool when a researcher wants to know whether readers are sensitive to any given manipulation, but it is limited in what it can reveal about how and when that manipulation would affect processing in more naturalistic reading scenarios. Although eye-tracking while reading may be a more appropriate method when the research question involves the latter, the Maze task could still prove useful in this type of situation as a way to test whether experimental materials are working properly before conducting an eye-tracking while reading study, which typically requires more time and resources.

The results from Experiment 2 also provide reason to prefer the Maze to SPR. The SPR results in Experiment 2 differed from the Maze in Experiment 1 in that the interaction of POS and STRESS was found *after* the point of disambiguation. As in previous comparisons of Maze and SPR focused on syntactic processing phenomena, it appears that implicit prosodic effects are localized in the Maze, but appear downstream in SPR. Since effects in a spillover region can be difficult to interpret, and because effects that are distributed across multiple words may be difficult to detect, the localized results in the Maze present a clear advantage over SPR. For this particular study, the effect size was much greater in the Maze experiment, which might also suggest that the Maze would be more robust to noisier data.

2.6 Implicit phrasal prosody

The results presented thus far have shown that both the Maze and SPR are sensitive to metrical prosody, despite concerns that the artificial nature of these tasks could interfere with the assignment of implicit prosody. This is promising for the study of implicit prosody, and suggests that both the Maze and SPR could be useful to researchers in this area, in particular when eye-tracking is unavailable or saved for a later stage of research. However, the conclusions I am licensed to draw about the viability of these tasks is limited because these experiments only addressed the sensitivity of these tasks to lexical stress. I would also like to establish that these tasks are sensitive to phrasal prosody, in particular because this aspect of prosody could be particularly susceptible to task effects: as discussed in Section 2.2.1, having to pause at each choice point in the Maze could make it difficult to keep track of the overall prosodic contour, or could encourage the reader to assign more prosodic boundaries than they would in more typical reading situations.

The goal of the next four experiments is to (i) use the NP/Z garden path to show that the Maze does not encourage a boundary-heavy prosody, and (ii) replicate the finding that longer relative clauses are more likely to attach high, to show that well-established implicit prosodic effects also arise in the Maze and SPR. Together, these replications suggest that the Maze is mostly comparable to previous studies on implicit prosody, which, given that it also has more localized effects than SPR, favors the use of the Maze when an alternative to eye-tracking while reading is needed.

2.7 Experiment 3: NP/Z in the Maze

As discussed in Section 2.2.1, one way the Maze could interfere with implicit prosody is by encouraging readers to assign a staccato, boundary-heavy prosody, because the task causes readers to take a pause at each word in the sentence to consider the target and the foil and decide which one is a better continuation of the sentence. If the Maze introduced such a boundary-heavy prosody, this would mean that task demands affect prosodification in a way that differs from what readers would do in more typical reading situations, raising questions about this task's ability to address questions about default implicit prosody and its impact on parsing. As mentioned earlier, such a boundary-heavy prosody could potentially explain what Witzel et al. (2012) call "hyper-incrementality," which is the possibility that the Maze "may encourage readers to close the current clause or constituent whenever possible."

However, I also raised doubts about hyper-incrementality: Witzel et al. (2012) suggested hyper-incrementality because they did not find a penalty when an ambiguous NP/S coordination was disambiguated to the S parse. However, they did not establish that the Maze also introduced a penalty for the NP parse, as would be predicted under the hyper-incrementality account, and the S parse penalty has been found in later work on the Maze (Boyce et al., 2020). It is therefore unclear to what extent the Maze actually encourages hyper-incrementality.

Because the existence of hyper-incrementality would bolster concerns about the Maze enforcing boundary-heavy prosody, I tested whether the NP/Z ambiguity would also cause a garden path in the Maze. The NP/Z ambiguity was briefly discussed in Chapter 1, and the example in (1) is repeated in (32). In this structure, the NP *the house* could be either the object of the verb *leaves*, as in (32a), or the subject of a new clause, as in (32b). Following Late Closure, the parser preferentially analyzes the ambiguous NP as the direct object of the subordinate clause verb, resulting in processing difficulty for the Z parse (Frazier & Fodor, 1978; Frazier & Rayner, 1982), but the difficulty associated with the Z parse goes away when a prosodic boundary is placed after the subordinate clause verb (Kjelgaard & Speer, 1999).

(32) When Roger leaves the house...

  a. ...it's dark. (NP parse)

  b. ...is dark. (Z parse)

Like the NP/S coordination ambiguity, the NP/Z garden path involves an ambiguity about whether an ambiguous NP is part of the clause currently being processed or is the subject of a new clause, it can be disambiguated through punctuation, and the competing parses have different prosodies. Thus, this is an appropriate test of the generalizability of hyper-incrementality. Moreover, even if one is skeptical of the claim that hyper-incrementality is related to implicit prosody, showing that the Maze is capable of detecting closure commitments would still address an issue flagged by Witzel et al. (2012) as meriting further research. Finally, to my knowledge the NP/Z garden path has not been studied in the Maze before, so comparing this effect in the Maze and SPR is a useful extension of existing methodological comparisons.

2.7.1 Methods

2.7.1.1 Materials

There were 48 experimental items that crossed the sentence structure (STRUCTURE: NP vs. Z) with the presence of a comma at the end of the subordinate clause (PUNCTUATION: Comma vs. No Comma). An example sentence is provided in Table 2.5. These items were adapted from previous work on the NP/Z garden path (Anderson & Carlson, 2010; Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Frazier, Carminati, Cook, Majewski, & Rayner, 2006; Frazier & Rayner, 1982; Staub, 2007a, 2007b), and standardized so that the ambiguous NP always contained a determiner, an adjective, and a noun, and the matrix subject in the NP condition contained a determiner and a noun. The matrix verb was always followed by an NP (Det + N), a PP (Prep + Det + N), or both; these possibilities were evenly distributed across all items, resulting in 16 of each type. All NP/Z items are provided in Appendix C.

Table 2.5: Sample NP/Z Item

STRUCTURE	PUNCTUATION	Item
NP	Comma	As the audience booed the unfunny comedian, the manager left the club in a hurry.
	No Comma	As the audience booed the unfunny comedian the manager left the club in a hurry.
Z	Comma	As the audience booed the unfunny comedian left the club in a hurry.
	No Comma	As the audience booed, the unfunny comedian left the club in a hurry.

If the Maze does not introduce hyper-incrementality or an atypical, boundary-heavy prosody, then I would expect the usual preference for the NP parse to come out in the Maze: upon encountering the ambiguous NP, participants would still be guided by Late Closure and/or by default implicit prosody to group this NP with the preceding material and attach it as the direct object of the subordinate clause verb. This garden path would be avoided in the Z, Comma condition, where the presence of a comma immediately alerts participants to the fact that there is no direct object, before they even see the ambiguous NP. However, they would incorrectly take the garden path in the Z, No Comma condition, and would not get disconfirming evidence until the matrix verb (*left*). Thus, under this view, I expect the classic pattern of slower response times in the Z, No Comma condition compared to the Z, Comma condition at the matrix verb, reflecting the need for reanalysis. The need to complete reanalysis could also lead to more Maze errors and/or worse comprehension question performance in the Z, No Comma condition.

If instead the Maze introduces "hyper-incrementality" and encourages the parser to close clausal constituents whenever possible, as proposed by Witzel et al. (2012), then upon encountering an optionally intransitive verb like *booed*, the parser should take advantage of the opportunity to end the subordinate clause at this position and adopt the intransitive reading of the verb (*booed*) (i.e., the Z parse). This would lead the parser to interpret the ambiguous NP (*the unfunny comedian*) as the subject of the matrix clause. This would circumvent the traditional garden path, such that there would be no penalty for the Z, No Comma condition relative to the Z, Comma condition, because hyper-incrementality would lead the parser to pursue the Z parse even in the absence of a comma. However, this could lead to a "reverse" garden path, with slower response times in the NP condition, once the parser has evidence that this ambiguous NP actually should have been analyzed as the direct object of *booed*. This reverse garden path could appear in at least two places. At the last word of the ambiguous NP (*comedian*), punctuation would disconfirm the expected Z parse in the NP, Comma condition but not in the NP, No Comma condition; this should result in a slowdown at this word in the NP, Comma condition relative to the NP, No Comma condition, reflecting the reanalysis from the Z parse to the NP parse. For the NP, No Comma condition, the Z parse would not be disconfirmed until the following word, the determiner of the matrix clause subject (*the manager*): the parser would realize that this noun is the actual subject of the matrix clause, and that it should have included *the unfunny comedian* in the subordinate clause rather than closing this clause after the verb. Thus, there could be longer reading times for the NP, No Comma condition relative to the NP, Comma condition at the determiner, reflecting the fact that a hyper-incremental parser would need to revise from the Z parse to the NP parse at this position.

Similarly, if the Maze introduces boundary-heavy implicit prosody, then participants would assign more implicit prosodic boundaries, even in places where such boundaries are optional rather than required. Under this view, the parser should place an implicit prosodic boundary after the subordinate verb *booed*, and then interpret this prosodic boundary as reflecting syntactic constituency, leading the parser to also close the subordinate clause at this position (i.e., pursue the Z parse). Again, the parser would interpret the ambiguous NP as the subject of the matrix clause, which would then lead to a "reverse" garden path in the NP condition when it turns out that it was actually the object of the subordinate clause verb. The pattern of results would be similar to the one described in the previous paragraph for hyper-incrementality. As noted above, hyper-incrementality and the hypothesized boundary-heavy prosody are not mutually exclusive, and the latter could in fact be the mechanism by which hyper-incrementality would arise.

2.7.1.2 Participants

This experiment was run with the homographs study in Experiment 1. As described in Experiment 1, 196 native English speakers participated in the study. After excluding participants who successfully completed less than 50% of the sentences in the Maze, data from a total of 180 participants was submitted to the analysis: 60 participants from Prolific Academic, who were run on the version with Mostly Stress Matched foils, and 120 undergraduates from UC Santa Cruz, 60 of whom were run on the version with Stress Mismatched foils and 60 of whom completed the version with Stress Matched foils. Because the foil manipulation was already addressed in Experiment 1, and because the stress pattern of the foils was not expected to (and did not) have an impact on the NP/Z garden path, this manipulation will not be discussed further.

2.7.1.3 Procedure

The same procedure described in Experiment 1 was used.

2.7.1.4 Response times

Extreme response times less than 100ms or longer than 5000ms were removed from the analysis. This resulted in a loss of .4% of the data. The response times at the subordinate clause verb through the end of the sentence are plotted in Figure 2.11. Note that this plot does not include the portion of the subordinate clause that preceded the verb, which was always between two to five words long. Additionally, because the items varied in whether the matrix verb (*left*) was followed by an NP, a PP, or both, the spillover region after the matrix verb varied in length between two to five words; there are therefore fewer observations at the third, fourth, and fifth spillover words (*in a hurry*). Impressionistically, the response times clearly reflect the classic NP/Z garden path: at the matrix verb, there is a slowdown in the Z, No Comma condition compared to the Z, Comma condition.

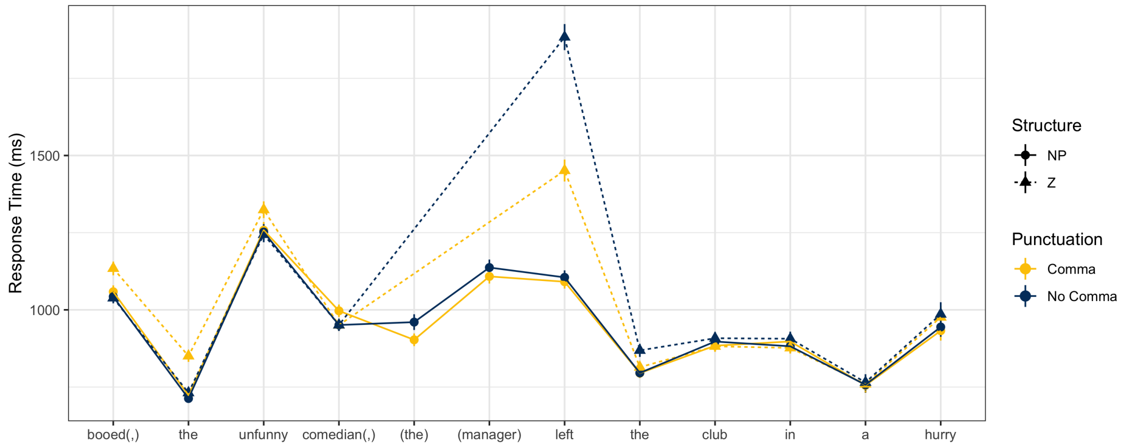


Figure 2.11: Mean response times by word in Experiment 3. Error bars show a 95% confidence interval.

Table 2.6 reports the mean RTs and standard error at the matrix verb and the first three spillover words. Bayesian mixed effects models were fit to the response times at each of these four words, with STRUCTURE, PUNCTUATION, and their interaction as fixed effects and the maximal random effects structure.

At both the matrix verb and the first spillover word, there was a main effect of STRUCURE (Matrix Verb: 589, [521, 657], Spillover 1: 47, [24, 71]), such that there was a slowdown in the Z condition, a main effect of PUNCTUATION (Matrix Verb: 233, [186, 278], Spillover 1: 26, [9, 43]), such that there was a slowdown in the No Comma condition, and a significant interaction (Matrix Verb: 436, [345, 525], Spillover 1: 56, [19, 94]), such that the No Comma slowdown was longer in the Z condition. There were no significant effects or interaction at the second spillover word (STRUCTURE: 4, [-27, 34], PUNCTUATION: 19, [-1, 38], Interaction: 21, [-15, 57]). At the third spillover word, the significant interaction reemerged (46, [7, 86]), but there was no main effect of STRUCTURE (-4, [-28, 21]), nor a main effect of PUNCTUATION (5, [-18, 28]).

STRUCTURE	PUNCTUATION	Matrix Verb	Spillover 1	Spillover 2	Spillover 3
NP	Comma	1091 (11)	795 (8)	884 (9)	898 (11)
	No Comma	1105 (12)	796 (8)	897 (10)	882 (10)
Z	Comma	1451 (18)	814 (9)	882 (9)	877 (9)
	No Comma	1884 (22)	869 (11)	908 (10)	906 (12)

Table 2.6: Mean Response Time (ms) at Matrix Verb and Spillover Region in Experiment 3

2.7.1.5 Maze error rates

Figure 2.12 shows the mean error rates by condition at each word from the subordinate clause verb to the end of the sentence. There is a clear spike in errors at the matrix verb *left*, particularly in the Z, No Comma condition, which parallels the response time data.

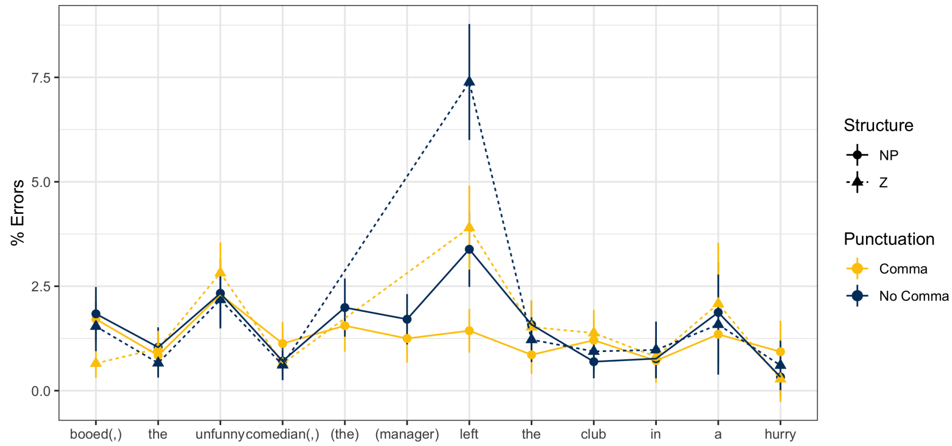


Figure 2.12: % of Maze errors by position in Experiment 3. Errors bars show a 95% confidence interval.

A Bayesian logistic mixed effects model of the probability of committing an error was fit to the accuracy data at the matrix verb, with STRUCTURE, PUNCTUATION, and their interaction as fixed effects and the maximal random effects structure. There was a main effect of STRUCTURE (1.3, [.73, 1.91]), reflecting a higher error rate in the Z condition, and a main effect of PUNCTUATION (.86, [.46, 1.28]), reflecting a higher rate in the No Comma condition, but no significant interaction (.42, [-1.18, .47]).

2.7.1.6 Comprehension question accuracy

Comprehension question accuracy was generally high, as summarized in Table 2.7. A Bayesian logistic mixed effects model was fit to the question data with the same effects structure as previous models. There was a main effect of STRUCTURE

(1.23, [.52, 1.95]), such that participants were more accurate in the Z condition, and a significant interaction (-.7, [-1.23, -.19]), such that the absence of a comma led to worse performance in the Z condition.

STRUCTURE	PUNCTUATION	% Correct
NP	Comma	87 (1.0)
	No Comma	87.8 (.9)
Z	Comma	94.3 (.6)
	No Comma	91.9 (.7)

Table 2.7: % correct responses to comprehension questions in Experiment 3

2.7.2 Discussion

This experiment clearly replicated the NP/Z garden path in the Maze: there was a slowdown in response times at the disambiguating matrix verb in the Z condition, and the interaction of STRUCTURE and PUNCTUATION shows that this slowdown was of a greater magnitude when there was no comma to guide the reader to the Z parse. Moreover, I found that participants had higher error rates on the Maze task in the Z condition, as reflected by the main effect of STRUCTURE. Finally, the interaction of STRUCTURE and PUNCTUATION in the comprehension questions showed that participants were less accurate in responding to questions in the Z, No

Comma condition when punctuation failed to guide them to the Z parse. Together, these findings show that participants expected the NP parse and experienced processing difficulty in the Z, No Comma condition when there was no punctuation to prevent the garden path. Since the Maze did not neutralize or reverse the bias for the NP parse, I conclude that the task does not introduce hyper-incrementality or a boundary-heavy prosody, at least not to the extent that they can override syntactic parsing principles like Late Closure.

There were also a few unexpected findings, including a STRUCTURE penalty at the matrix verb such that there was a slowdown in the Z condition even when the comma should have prevented a garden path, and the finding that participants were overall more accurate when responding to questions in the Z condition with the garden path. I set these findings aside for now and return to them in the discussion of Experiment 4.

One may ask why this study found the usual NP/Z garden path, yet Witzel et al. (2012) did not find the NP vs. S coordination garden path, which also involves clausal closure. As noted above, Boyce et al. (2020) replicated the NP/S garden path in the Maze, which suggests that the original null result may have been due to the study being underpowered, rather than being due to a difference between the coordination ambiguity and the NP/Z garden path. In general, I contend that the present results converge with Boyce et al. (2020) and suggest that the Maze may not be as hyper-incremental as previously thought. Crucially, I also interpret the presence of the garden path as showing that participants did not construct a boundary-heavy prosody, which would have encouraged them to prefer the Z parse and led to processing difficulty in the NP condition rather than the Z condition.

Alternative interpretations are available, however. First, it is possible that the Maze *did* encourage a boundary-heavy prosody, but this prosody was not strong enough to override other factors that lead to the preference of the NP parse, such as Late Closure and the likelihood that each verb should take a direct object. While I cannot directly rule out this possibility, previous studies using auditory stimuli have shown that the presence of a prosodic boundary after the subordinate clause verb erases the penalty associated with the Z structure (Kjelgaard & Speer, 1999); thus, I find it unlikely that readers would insert a boundary in this position yet still be susceptible to the garden path, and find it simpler to assume that the Maze did not introduce extra implicit boundaries.

In light of this alternative, a more secure claim is that I failed to find evidence for an atypical boundary-heavy prosody, but I have not yet shown evidence that the prosody being assigned is typical. In order to argue for the latter, I would need to detect the *presence* of an effect that is attributed to implicit prosody; showing that such an effect appears in both the Maze and other reading tasks would support the conclusion that the prosody in the Maze is similar to the prosody assigned in other tasks. I set this aside and return to the issue in Experiments 5 and 6, where I replicate the effect of relative clause length on attachment preferences.

2.8 Experiment 4: NP/Z in SPR

Having established the NP/Z garden path in the Maze, I now replicate this effect in self-paced reading. Unlike the Maze, there was no expectation that SPR would introduce a boundary-heavy prosody. If anything, the strategy of tapping the keyboard quickly, which was reported by many subjects in the debriefing, could encourage the opposite effect: a prosodic contour with fewer boundaries than in more natural reading settings. In this case, any effect of SPR-specific prosody would simply reinforce the extant preference for the NP parse, leading to the usual garden path. I therefore do not speculate further on the impact of SPR on implicit prosody in the discussion of this experiment. Instead, the replication serves as a sanity check that the items are behaving as expected, granting additional confidence in the results from Experiment 3. In addition, because this chapter is focused on methods, conducting the NP/Z experiment in SPR allows for further cross-methodological comparisons.

2.8.1 Methods

2.8.1.1 Materials

The same experimental items from Experiment 3 were used, sans foils.

2.8.1.2 Participants

Experiment 4 was run as part of the same study as Experiment 2, so the same 65 undergraduate native English speakers at UC Santa Cruz participated for course credit. As in Experiment 2, five participants who participated after the target of 60 subjects had been reached were excluded.

2.8.1.3 Procedure

The same procedure described in Experiment 2 was used.

2.8.2 Results

2.8.2.1 Response times

As in the previous Maze and SPR experiments, extreme observations less than 100ms or greater than 5000ms were eliminated from the data prior to analysis, resulting in a loss of .2% of observations from the NP/Z sentences, including a loss of .2% of data from the matrix verb and the first three spillover words. The response times for the subordinate clause verb through the end of the sentence are plotted in

Figure 2.13. There is a clear slowdown in the Z, No Comma condition at the matrix verb, as in Experiment 3, but this penalty also appears to continue through the next few spillover words, in a departure from the Maze results.

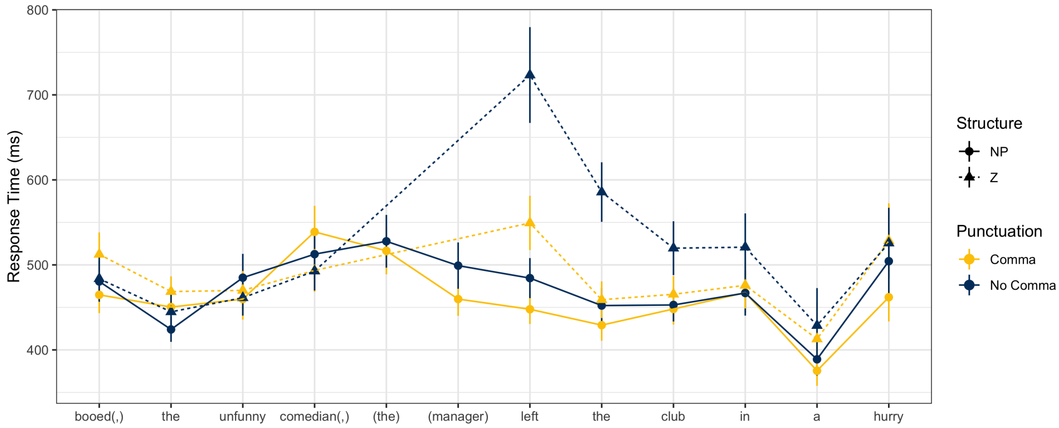


Figure 2.13: Mean response times by word in Experiment 4. Error bars show a 95% confidence interval.

The mean RTs and standard error at the matrix verb (*left*) and the first three spillover words (*the club in*) are reported in Table 2.8. Separate Bayesian mixed effects models were fit to the response times at the matrix verb and the three following spillover words, with STRUCTURE, PUNCTUATION, and their interaction as fixed effects and the maximal random effects structure. At the matrix verb (*left*), there was a main effect of STRUCTURE (171, [106, 233]), indicating longer RTs in the Z condition, a main effect of PUNCTUATION (107, [54, 159]), indicating longer RTs in the No Comma condition, and a significant interaction (140, [41, 239]), indicating that the processing cost incurred by the lack of a comma was stronger in the Z condition. At the first spillover word, *the*, there were the same effects of STRUCTURE (81, [46, 117]), PUNCTUATION (75, [43, 106]), and their interaction (104, [47, 160]). The same effects were found at the second spillover word, *club*: STRUCTURE (42, [15, 68]), PUNCTUATION (30, [3, 57]), and their interaction (50, [4, 98]). At the third spillover word, *in*, there was no significant effect of STRUCTURE (29, [-5, 64]), PUNCTUATION (26, [-8, 58]), nor a significant interaction (48, [-17, 113]).

STRUCTURE	PUNCTUATION	Matrix Verb	Spillover 1	Spillover 2	Spillover 3
NP	Comma	448 (9)	429 (9)	448 (9)	467 (12)
	No Comma	484 (12)	452 (11)	453 (10)	467 (14)
Z	Comma	549 (16)	459 (11)	465 (11)	476 (14)
	No Comma	723 (29)	586 (18)	520 (16)	521 (20)

Table 2.8: Mean response times (ms) at matrix verb and spillover region in Experiment 4.

2.8.2.2 Comprehension question accuracy

The overall accuracy across conditions is summarized in Table 2.9, with standard error computed over participant means. As in Experiment 1, performance on the comprehension questions was high overall. A Bayesian logistic mixed effects model was fit to the comprehension data with the STRUCTURE, PUNCTUATION, and their interaction as fixed effects, and the maximal random effects structure.

There was a main effect of STRUCTURE (.96, [.22, 1.73]), such that participants were more likely to answer correctly in the Z condition. There was no main effect of PUNCTUATION (-.05, [-.43, .35]) or a significant interaction (-.36, [-1.14, .38]).

STRUCTURE	PUNCTUATION	% Correct
NP	Comma	86.1 (1.6)
	No Comma	86.8 (1.5)
Z	Comma	92.9 (.8)
	No Comma	91.4 (1.2)

Table 2.9: % correct responses to comprehension questions in Experiment 4

2.8.3 Discussion

As in Experiment 3, the NP/Z garden path was replicated, as reflected in response times at the matrix verb and the first two spillover words. This is encouraging, because it reflects the findings from the Maze task; moreover, there was no reason to expect the garden path *not* to appear in SPR, so any alternative results would have been concerning.

The Maze and SPR also shared some unexpected findings. In the response times at the matrix verb, the main effect of STRUCTURE showed that there was a slowdown in the Z condition, even when the presence of a comma should have prevented the the garden path. One possibility is due to the relative position of the matrix verb in the different conditions: in terms of relative position, the verb appears two words later in the NP condition compared to the Z condition due to the addition of the matrix subject (*the manager*). If participants read faster as they progress through the experiment, then the later position of the matrix verb in the NP condition could explain why the RTs are faster in this condition. Alternatively, it is possible that participants are aware of the fact that commas sometimes appear and sometimes do not; due to this unreliability, they might not attend as closely to the commas over the course of the experiment. If participants who expect commas to be unreliable begin to ignore or downweight punctuation at some point, then they might not be as strong of a cue that participants are in the Z parse, which could lead to them still being (somewhat) garden pathed in the Z, Comma condition, explaining why there is still a slowdown despite the disambiguating punctuation. Another possibility is that even with disambiguation, there is something dispreferred about the Z parse that makes it harder to process. For instance, perhaps the relative infrequency of this construction makes it more difficult to parse, or the intransitive reading of the subordinate clause verb was dispreferred even with the comma, such that this affected processing downstream. To my knowledge, most previous work on the NP/Z garden path has not crossed STRUCTURE and PUNCTUATION as I did in Experiments 3 and 4, and typically contrast either the Z, Comma and the Z, No Comma conditions, or the Z, No Comma and the NP, No Comma conditions (Christianson et al., 2001; Frazier, Carminati, et al., 2006; Frazier & Rayner, 1982; Staub, 2007b). As such, they do not shed light on whether a slowdown at the matrix verb in the Z, Comma conditions relative to the NP conditions is expected. At the very least, I can conclude that the slowdown in the Z, Comma condition is not Maze-specific, because it also showed up in Experiment 4.

Additionally, in both experiments the question data showed participants were more accurate in the Z condition than in the NP condition; performance was also worse for the Z, No Comma condition relative to the Z, Comma condition in the Maze but not in SPR. Since the Z parse was dispreferred, one might have expected comprehension accuracy to be worse for the Z condition relative to the NP condition, under the assumption that the garden path would interfere with participants' understanding of the sentence. However, the Z conditions were also shorter on average, and it is plausible that this shorter length made it easier to remember the sentences when responding to the questions. Additionally, because the Z parse is generally harder to process, it could be the case that failure on the Maze task filtered out the particularly difficult trials, such that the trials analyzed in the comprehension data are those on which participants understood the sentence better, leading to higher comprehension performance than otherwise expected. This last explanation could not explain the better performance in the Z condition in SPR, however.

There were also several differences between Experiments 3 and 4, some of which reflect previously documented differences between the Maze and SPR. For instance, the effect size was smaller in SPR, and there was a slight difference in timing. While the bulk of the effect occurred at the disambiguating verb in the Maze, lingering into the first spillover word, the effect was more evenly distributed across the verb and the first two spillover words in the SPR experiment. Overall, this confirms that while both the Maze and SPR are sensitive to similar effects, there are reasons to prefer the Maze when possible.

2.9 Experiment 5: Size Effects on RC Attachment in the Maze

The previous two experiments have shown that the NP/Z garden path exists in both the Maze and SPR, providing suggestive evidence that the Maze does not introduce a boundary-heavy prosody, which would have reversed the NP/Z garden path. However, as pointed out in the discussion of Experiment 3, the argument that the Maze does not lead to an atypical implicit prosody would be bolstered by further replications showing that the Maze is also sensitive to effects that are *caused* by a particular prosody. To that end, I focus on Fodor's (2002a, 2002b) finding that relative clause length affects attachment preferences in sentences like those in

(33). In this sentence, the attachment site of the relative clause is ambiguous: it could be describing *the mentor* or *the writer*. However, as described in Chapter 1, a long line of work has shown that the length of the relative clause affects readers' interpretation: readers are more likely to provide a high attachment interpretation, saying that the RC modifies the first noun *the mentor*, when the RC is longer, as in (33b). Again, as described earlier, this finding has been influential in the development of the Implicit Prosody Hypothesis, as Fodor (2002a, 2002b) suggested that the effect of length is due to the preference for a prosodic break before a long RC, which is then interpreted by the parser as reflecting syntactic constituency. This effect of a pre-RC boundary on attachment has been documented in various listening studies (Bishop et al., 2015; Fromont et al., 2017; Maynell, 1999), supporting Fodor's claim that an implicit boundary in this position should encourage non-local attachment.

(33) a. The mentor of the writer [who speaks limited French] advised never to go with a hippie to a second location.

b. The mentor of the writer [who speaks limited French and hasn't cried once today] advised never to go with a hippie to a second location.

Under the assumption that RC length effects are due to implicit prosody, then establishing that length has the same effect on attachment preferences in the Maze would support the conclusion that the implicit prosody assigned in the Maze is sufficiently similar to the implicit prosody assigned in more natural reading settings, because it would be necessary for the length effect to encourage prosodic boundaries in the same location in order for the effect on attachment preferences to be the same across tasks.

To test whether the Maze is sensitive to RC length effects, I replicated part of Hemforth et al. (2015). In their study, they conducted a series of questionnaires in English, Spanish, German, and French, showing that across languages, longer RCs were more likely to lead to high attachment preferences, thereby replicating the classic Fodor effect. Since the present study is focused on the sensitivity of the Maze to length effects, and I was primarily interested in using their items because they are known to exhibit the length effect, I only conducted a Maze and SPR study using their English items. However, it should be noted that they were interested in cross-linguistic differences in RC attachment and how these also related to whether the complex NP was in an object or non-object position; since this is not relevant to the replication of the length effect in English, I do not discuss the language manipulation further except where helpful to clarify how the present findings relate to theirs.

2.9.1 Methods

2.9.1.1 Materials

The 40 experimental items consisted of the 32 items from Hemforth et al. (2015) and eight additional items, for a maximum possible total of 10 observations per condition per participant. Items were distributed across four lists via Latin Square. An example item is provided in Table 2.10. Following the original experiment, there were two factors: POSITION, which referred to whether the complex NP was in the Subject[2] or Object position, and SIZE, which referred to whether the RC was Short or Long.

Position	Size	Sentence			
		Preamble	Complex NP	RC	Post-RC
Subject	Short		The son of the colonel	who tragically died of a stroke	wrote five books on tropical disease.
	Long		The son of the colonel	who tragically died of a stroke	wrote five books on tropical disease.
Object	Short	The doctor met	the son of the colonel	who died.	
	Long	The doctor met	the son of the colonel	who tragically died of a stroke.	

Table 2.10: Example item from Experiments 5 and 6

Although the sentences were presented one word at a time in the experiment, the sentence was divided into four regions for the analysis of the response times, because the sentences varied in length; these regions are exemplified in Table 2.10.

[2]Hemforth et al. (2015) refer to this as "non-object" position, in order to reflect that in Spanish and German, which were also investigated in their study, there are multiple preverbal positions available depending on whether the subject is a topic. Since the focus here is on English, I adopt the term "subject position" for expository ease.

The Preamble region, which consisted of the matrix subject and verb that preceded the complex NP in the Object condition, was always two or three words long. The Complex NP was always five words long, and was the only region that never varied in length across items. The RC was always two words long in the Short condition, and varied in length between four to six words in the Long condition. The Post-RC region, which followed the RC in the Subject condition and consisted of the matrix verb and any following material, was always between three to eight words long.

Each sentence was followed by a comprehension question that probed whether participants interpreted the relative clause as applying to the first or second noun in the complex NP. The question differed between the Short and Long conditions because it always repeated the entire RC from the preceding sentence, as illustrated in 2.11. All sentences are provided in Appendix D.

Size	Question
Short	Who died?
Long	Who tragically died of a stroke?

Table 2.11: Example comprehension question from Experiments 5 and 6

Based on the results from the preceding experiments, the Maze is not expected to interfere with the assignment of implicit prosody. I therefore predict that Hemforth et al.'s (2015) main effect of SIZE will be replicated, such that participants will be more likely to provide high attachment interpretations when the RC is long compared to when the RC is short, because the same pressure for prosodic balance that, by hypothesis, favored high attachment of the long RC in the original study should also be present in the Maze task. There may also be an interaction of SIZE and POSITION, such that the effect of SIZE is stronger when the complex NP is in Object position, because Hemforth et al. found this trend in their data, although the interaction was not significant.

In contrast, if the Maze interferes with the assignment of implicit prosody and/or introduces a boundary heavy prosody, then an effect of SIZE on attachment preferences would not be expected. If readers (subconsciously) recognize that phrasing is disrupted and/or shaped by the task forcing them to pause at each word, rather than being informative about the structure of the sentence, then the parser could downweight the boundaries and ignore them. In this case, no effect of SIZE would emerge: even if a different prosodic phrasing emerged between conditions, this phrasing would not influence syntactic parsing. Another possibility is that readers would still attend to the implicit prosodic boundaries and use them to guide syntactic parsing, but that having more frequent boundaries would result in a qualitatively different contour that impacts parsing differently than in other reading tasks. For instance, it is possible that a boundary-heavy prosody would encourage readers to construct a prosodic boundary between the complex NP and the RC in all cases, and not only when the RC is long. If the presence of the boundary separating the complex NP from the RC, which is hypothesized to drive the size effect in previous experiments, no longer depends on the RC length but instead is always present, then RC length would no longer be expected to have an effect on the rate of high attachment interpretations.

2.9.1.2 Participants

66 native English speakers participated in the experiment. 35 subjects were undergraduates at UC Santa Cruz and received course credit for participating, and 31 subjects were recruited from Prolific Academic and were paid $8. Six subjects were excluded for failing to successfully complete the Maze for at least 50% of the sentences, leaving data from 60 subjects for the analysis.

2.9.1.3 Procedure

The same procedure from Experiments 1 and 3 was used. For the experimental items, the two response options for the comprehension questions were the two nouns in the Complex NP structure; participants pressed the "e" key to choose the answer on the left and the "i" key to choose the answer on the right. For filler items, the options were always "Yes" and "No," and participants pressed the "e" key for "Yes" and the "i" key for "No."

2.9.2 Results

2.9.2.1 Comprehension question responses

Only trials on which participants successfully completed the Maze for the entire sentence were included in the analysis of the comprehension question responses; this left 1890 observations, or 79% of the 2400 trials, for analysis. The percentage of high attachment interpretations by condition is provided in Table 2.12. A Bayesian logistic mixed effects model was fit to the response data with POSITION, SIZE, and their interaction as fixed effects, and the maximal random effects structure.

There was no effect of POSITION (.24, [-.03, .52]). There was a main effect of SIZE (.33 [.08, .57]), such that participants were more likely to provide a high attachment interpretation in the Long condition, and an interaction (-.6, [-1.12, -.08]), such that this SIZE effect was neutralized when the complex NP was in subject position.

POSITION	SIZE	% High Attachment
Subject	Short	35.2 (3.5)
	Long	31.6 (3.1)
Object	Short	24.5 (2.5)
	Long	35 (2.6)

Table 2.12: % high attachment responses in Experiment 5

2.9.2.2 Maze error rates

Mean Maze error rates in each condition are shown in Figure 2.14. Overall, the error rates were relatively low, indicating that participants were able to successfully complete the task for the critical items.

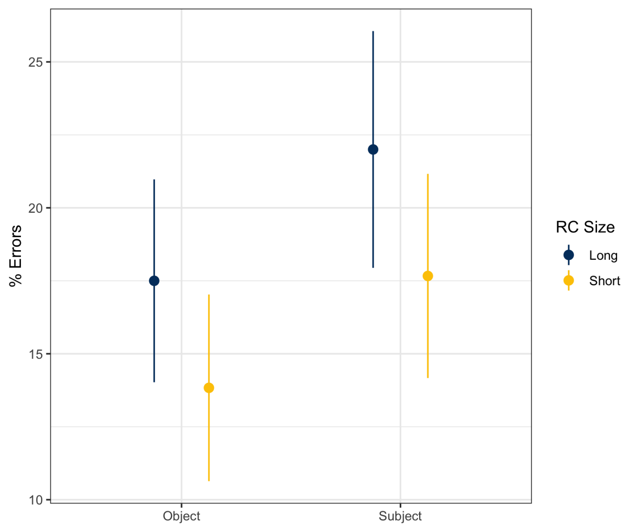


Figure 2.14: Mean Maze error rates by condition in Experiment 5. Error bars show a 95% confidence interval.

A Bayesian logistic mixed effects model was fit to the error data with POSITION, SIZE, and their interaction as fixed effects, and the maximal random effects structure.

There was a main effect of POSITION (.31, [.02, .6]), such that participants were more likely to make errors when the complex NP was in subject position. This was somewhat surprising, as I did not expect to find any differences in participants' ability to complete the Maze based on whether the complex NP was in subject or object position, because the same RC attachment ambiguity was present. However, one reason for this effect could be that the sentences in the Subject condition condition were on average slightly longer than those in the Object condition (Subject: 12.4 words vs. Object: 11.3 words). This means that there were more opportunities to make mistakes in the Subject condition, potentially explaining why there was a slightly higher error rate.

However, this explanation would suggest that SIZE should also be relevant for failure rates, with a higher error rate in the Long condition, because the Long condition was on average longer by definition (Short: 10.3 words vs. Long: 13.3 words). Although numerically the percentages do suggest greater failure rates in the Long condition, there was no main effect of SIZE (.23, [-.17, .62]) or a significant interaction (.04, [-.49, .58]). Thus, it is unclear if the main effect of POSITION is spurious or if it is driven by a different number of opportunities for errors across conditions or by some other factor (i.e., perhaps for whatever reason there are stronger constraints placed on possible continuations when the complex NP appears in object position, making the foils easier to reject and leading to fewer errors in this condition).

Since I did not have a reason to expect any difference between these conditions a priori, I set the main effect of POSITION aside and do not discuss it further. The main point of examining the error rates is to show that participants generally performed well in the task, and that it is important to consider whether different sentence lengths could be driving differences in error rates when interpreting this dependent variable.

2.9.2.3 Response times

Only response times for trials on which participants did not make any errors were included in the analysis. Response times for each region were calculated by taking the sum of the observations for each word in that region. I excluded a region if any of the individual observations within that region had a response time shorter than 100ms or greater than 5000ms; this excluded 1.3% of the data. The mean response times on each region are shown in Figure 2.15.

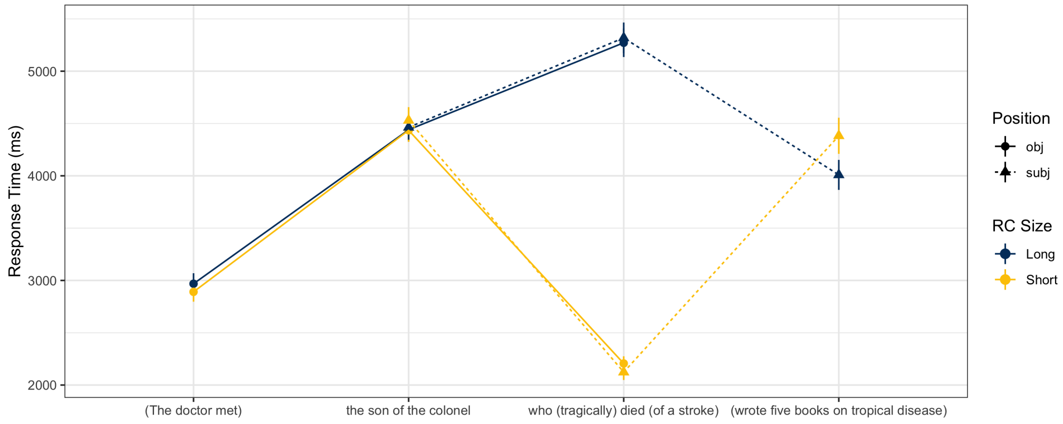


Figure 2.15: Response times by region in Experiment 5. Error bars show a 95% confidence interval.

Bayesian mixed effects models were fit to the response times at each region. For the response times at the Preamble and Post-RC regions, SIZE was the only fixed effect, because the Preamble was only seen in the Object condition, and the Post-RC region was only seen in the Subject condition. For the Complex NP and RC regions, the fixed effects were POSITION, SIZE, and their interaction. All models had the maximal random effects structure.

At the Preamble, I failed to find a main effect of SIZE (73, [-54, 203]). This finding is encouraging: I would not have expected any differences, since the words within the region were identical across conditions. Similarly, at the complex NP, I failed to find a main effect of POSITION (66, [-62, 197]), SIZE (-24, [-130, 82]), or a significant interaction (-84, [-296, 124]).

At the relative clause, there was an effect of SIZE (3168, [2912, 3428]) but no effect of POSITION (-30, [-132, 71]) or an interaction (98, [-64, 321]). Again, this is not surprising: the SIZE effect is clearly driven by the fact that the RC included several more words in the Long condition. Beyond this, I did not expect any processing differences between conditions that would result in a slowdown, and this is reflected in the lack of an effect of POSITION and the lack of an interaction.

At the Post-RC region, there was a main effect of SIZE (-329, [-491, -170]), such that response times were longer in the Short condition. Note that this region was identical between the Subject, Short and Subject, Long conditions, so this effect is not caused by a difference in material. I set this finding aside for now and address potential reasons for this effect in the discussion of Experiment 6.

2.9.3 Discussion

The primary goal of this experiment was to determine whether well-known effects of RC length on RC attachment, which have most frequently been investigated in studies that employ questionnaires in which entire sentences are visible at once, would also come out in the Maze. As described above, previous work has found that longer relative clauses are more likely to attach high compared to shorter relative clauses, and one popular explanation of this finding appeals to implicit prosody: a long RC prefers to attach high in order to achieve prosodic balance with the preceding constituent (Fodor, 2002a, 2002b; Hemforth et al., 2015).

I reasoned that if the implicit prosody assigned by readers in the Maze is the same as (or similar to) the implicit prosody assigned in more naturalistic reading situations, then I should replicate the findings of Hemforth et al. (2015), who showed the classic length effect on attachment. If instead the Maze interferes with the assignment of implicit prosody, or causes readers to adopt a boundary-heavy prosody distinct from what they would usually assign in silent reading, then phrase length should not have the same (or potentially, any) impact on attachment decisions in the Maze, because readers would no longer be constructing the same prosodic contour that, by hypothesis, causes long RCs to attach high.

Overall, the findings of this study replicate those of Hemforth et al. (2015). As in their study, participants were more likely to provide high attachment responses when the RC was long. The replication of this effect is promising, as the existence of a length effect in the Maze suggests that participants are able to construct an implicit prosody and that this prosody impacts parsing in a similar way as it does in more typical reading situations.

However, the results differed from theirs in a few ways. In the present study, the main effect of SIZE was qualified by an interaction of SIZE and POSITION showing that the effect of RC length on attachment preferences was limited to sentences in which the complex NP was in object position. In contrast, Hemforth et al. (2015) did not find a significant interaction of SIZE and POSITION in their study, though this interaction was trending toward significance, suggesting a weaker effect of SIZE when the complex NP was in subject position. However, it should be noted that the test for this interaction consisted of data from English, German, and Spanish, so it is not clear to what extent this trend reflects the pattern in English specifically. On the one hand, their English data numerically shows a smaller difference between the Short and Long conditions when the NP is in subject position, which would suggest that this trend is driven in part by the English data, similar to the interaction documented here. On the other hand, when they included language as a factor, there were no interactions involving language and SIZE, but there was a significant interaction of language and POSITION showing that the effect of POSITION was significant in German and Spanish but not English. These latter findings might suggest that the trending interaction of POSITION and SIZE in their study was due to the data from German and Spanish, not English. In that case, the interaction documented in the present study would be a divergence from their results.

Stepping back from the specifics of their results, it is clear that their English data show a numerical trend of a smaller SIZE effect when the complex NP is in subject position. This is somewhat consistent with the interaction of POSITION and SIZE in Experiment 5, with the exception that the present data suggest that the SIZE effect is not merely reduced but in fact neutralized when the complex NP is in subject position. Thus, one may consider this a partial replication of Hemforth et al. (2015), because I found that long RCs led to more high attachment responses and that this effect was strongest with the complex NP is in object position. Under the assumption that this effect is driven by implicit prosodic phrasing, this strongly suggests that readers are able to assign an implicit prosody in the Maze, and that this implicit prosody is fairly similar to the prosody they would assign in other reading tasks.

Less clear is why the SIZE effect should be neutralized (or potentially reversed) in the Subject condition. I consider it unlikely that this result is due to some property of the Maze task that interferes with implicit prosody, because I have now replicated effects of implicit prosody across multiple experiments in this chapter, including in the Object condition of the present study. It is unclear why readers would be able to assign an (apparently typical) implicit prosody in these other experiments and conditions yet not in the subject condition for this experiment. Moreover, though the absence of the SIZE effect with the NP in subject position prevents this from being a perfect replication of the original study, a weaker effect in subject position is exactly what I would expect based on Hemforth et al.'s (2015) results: thus, I contend that this slight difference in the overall pattern of results does not undermine my primary point, which is that implicit prosodic effects can be found in the Maze. Moreover, Hemforth et al. (2015) point out that most previous studies have investigated RC attachment with the complex NP in object position. Thus, while the effect of length is clearly established when the complex NP is in object position, more work on the effect of RC length when the complex NP is in subject position is needed in order to determine whether the findings for the Subject condition in the current study are exceptional or the norm.

In sum, based on the Maze experiment, I conclude that relative clause length has a similar effect on attachment in the Maze as it does in more natural reading tasks, validating the use of the Maze for experiments on implicit prosody. However, a skeptic of participants' ability to assign implicit prosody in the Maze could attempt to develop an alternative explanation: perhaps the effect of RC length on attachment height, in *any* task, is not due to implicit prosody, but instead is due to some as yet undiscovered factor. Under this view, the RC length effect was replicated in the Maze because this other factor is present in any reading task, and not because implicit prosody is present in the Maze. The challenge for such an account is to find a convincing non-prosodic reason why longer RCs favor high attachment. As discussed previously, many alternative explanations of RC attachment preferences have run into issues, while appeals to both overt and implicit prosody have allowed for a unified explain of variation in RC attachment preferences both within and across languages (Fodor, 2002a, 2002b). Thus, I am skeptical that a non-prosodic alternative would have the same empirical coverage as an account that attributes these effects to implicit prosody, and find it more straightforward to assume that the

Maze is sensitive to implicit prosody.

2.10 Experiment 6: Size Effects on RC Attachment in SPR

Having established that the effect of RC length on high attachment interpretations comes out in the Maze, I next ran the same materials in a self-paced reading study. This study had several purposes. First, I wanted to know whether implicit prosodic effects relating to phrasing and attachment preferences would also arise in self-paced reading. As discussed earlier, the SPR task may encourage a more superficial scanning strategy, by which participants tap quickly through the sentence in order to reach the comprehension question quickly before they have forgotten the content of the sentence. Such a strategy could lead to a situation in which the prosodic representation of a sentence is less detailed and/or lacking in as many boundaries as it would have when the sentence is presented all at once, as in a questionnaire task. This could lead to the absence of the SIZE effect if this superficial scanning causes participants to attend less to implicit prosody, and/or to be less likely to place a prosodic break between the complex NP and the RC, regardless of the size of the latter, because they are trying to reach the end of the sentence as fast as possible.

Second, this study extends the cross-methodological comparison of the Maze and SPR developed in Experiments 1 through 4, adding to our knowledge of when and on what measures these tasks converge and diverge. Since the Maze has been argued to have advantages over SPR, it is important to understand whether and how the nature of these tasks impacts implicit prosody, to see if the same advantages carry over to this area, or whether SPR should be preferred over Maze when eye-tracking is not available.

2.10.1 Methods

2.10.1.1 Materials

The same materials from Experiment 5 were used.

2.10.1.2 Participants

108 undergraduate native English speakers at UC Santa Cruz participated in the experiment for course credit.

2.10.1.3 Procedure

The same procedure from Experiments 2 and 4 was used.

2.10.2 Results

2.10.2.1 Comprehension question responses

The percentage of high attachment interpretations in each condition is summarized in Table 2.13. A Bayesian logistic mixed effects model was fit to the question response data with POSITION, SIZE, and their interaction as fixed effects, and the maximal random effects structure. There was a main effect of SIZE, such that participants were more likely to provide a high attachment interpretation in the Long condition (.36, [.17, .54]). There was no effect of POSITION (-.11, [-.38, .15]), nor was there a significant interaction (-.23, [-.57, .1]).

POSITION	SIZE	% High Attachment
Subject	Short	30.4 (2.6)
	Long	34.2 (2.7)
Object	Short	29.3 (2.2)
	Long	36.9 (2.2)

Table 2.13: % high attachment responses in Experiment 6

2.10.2.2 Response times

As in Experiment 5, I calculated the response time for each region and excluded from the analysis any region for which at least one observation was shorter than

100ms or greater than 5000ms; this excluded 2.3% of the data. The mean response times for each region are shown in Figure 2.16.

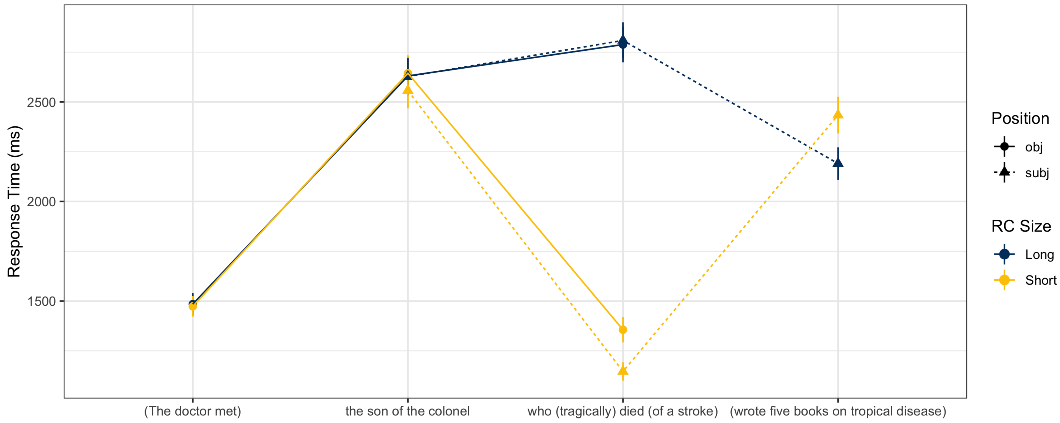


Figure 2.16: Mean response times by region in Experiment 6. Error bars show a 95% confidence interval.

I fit Bayesian mixed effects models to the response times at each region, with the same effects structure as in Experiment 5. At the Preamble region, there was no significant effect of SIZE (13, [-63, 90]). At the Complex NP, there was no effect of POSITION (-42 [-151, 64]), no effect of SIZE (14, [-51, 80]), and no significant interaction (64, [-74, 200]).

At the relative clause, there was a significant effect of POSITION (-97, [-172, -24]), such that the RC was read faster when the complex noun phrase was in subject position, a main effect of SIZE (1551, [1389, 1716]), such that the RC was read faster when it was shorter, and an interaction (237, [101, 372]), such that the speed-up in the short condition was greater when the complex NP was in subject position. The main effect of SIZE was expected, because this region consisted of more words in the Long condition. However, the effect of POSITION and the interaction were not expected. One possibility is that these findings reflect a wrap-up effect, such that response times are longer in the Object condition because the RC is located at the end of the sentence in this condition, although it is worth noting that this effect and interaction were not significant in the Maze study. I set these findings aside because I did not have any a priori expectations for this region, nor a reason to expect why these slight differences in response times would play a role in the interpretation responses, which are the primary measure of interest. Because this is a methods chapter, the motivation behind analyzing the response times was to show that results were comparable between conditions within each experiment, as well as comparable between the Maze and SPR: if there had been substantially different response times between tasks or conditions, then this would have been a red flag. Since outside of this RC region the response times are comparable across experiments and do not show great variation, it seems unlikely that differences in button pressing behavior between conditions and/or tasks, as reflected in the response times, is driving the documented effects on interpretation.

Finally, at the Post-RC region, there was a main effect of SIZE (-244, [-343, -146]), such that response times were longer in the Short condition. This replicates the finding from the Maze study, and I return to potential explanations for this effect in the discussion.

2.10.3 Discussion

The purpose of this experiment was to test whether effects of implicit prosodic boundaries that arise in questionnaires with whole-sentence presentation would also arise in SPR, where task strategies such as tapping quickly could potentially impact the assignment of implicit prosody. The results also provide a benchmark against which to compare the findings from the Maze experiment.

The study successfully replicated the main finding from Hemforth et al. (2015): there was a main effect of SIZE, such that participants were more likely to provide high attachment interpretations when the complex NP was followed by a long RC compared to when it was followed by a short RC. Under the view that this effect of length is due to prosodic balance and the insertion of an implicit prosodic boundary before a long RC, then this establishes that SPR is sensitive to effects of implicit prosodic phrasing, despite potential task parameters that could plausibly influence implicit prosody.

Unlike the Maze task, but like Hemforth et al. (2015), I did not find an interaction of POSITION and SIZE in this experiment. While this means that the SPR experiment more closely replicated the original results from Hemforth et al. (2015), I caution against the interpretation that the implicit prosody used in SPR is more similar to the implicit prosody assigned in situations with whole-sentence presentation compared to the implicit prosody in the Maze, which might suggest adopting SPR over the Maze in studies of implicit prosody. Even though the interaction was not significant, the difference between the Short and Long conditions was still numerically smaller in the Subject condition than in the Object condition in the SPR experiment, and this interaction was also trending in Hemforth et al. (2015). This suggests that the Maze-specific interaction is still representative of a trend that exists across all tasks, although it should be noted that the interaction in the Maze suggested that the SIZE effect was neutralized in the Subject condition, and not just smaller.

Finally, it is worth addressing the response time data. The response time data were primarily analyzed to serve as a sanity check, showing that response times to regions that were identical across conditions did not vary in unexpected ways. In general, this is what was found: at almost every region, there were either no differences when the regions were identical across conditions, or the differences were easily explained by the difference in the length of the region between conditions, as with the main effect of SIZE at the relative clause. However, in both the Maze and SPR, there was a main effect of SIZE in the Post-RC region. This region contained identical material in the Subject, Short and Subject, Long conditions, yet the response times were longer in the Subject, Short condition. The crucial difference is that the Post-RC region was immediately preceded by a longer RC in the Subject, Long condition. One possibility is that this is an Anti-Locality effect.

The term “Anti-Locality” refers to a class of findings in which increasing the distance between a head and its dependents *facilitates* processing. For instance, in Jaeger et al.’s (2005) self-paced reading study, cited by Levy (2008) and Vasishth and Drenhaus (2011), participants read sentences like those in Table 2.14 in which a matrix subject was modified by an object relative clause that ended with one, two, or three prepositional phrases before returning to the matrix clause to complete the subject-verb dependency. Jaeger et al. (2005) found that mean reading times at the matrix verb were significantly longer in the 1 PP condition compared to both the 2 PP and 3 PP conditions; there was no significant difference between the 2 PP and 3 PP conditions. In other words, increasing the distance between the matrix subject and the matrix verb by extending the length of the relative clause actually led to faster response times at the matrix verb.

	The player [that the coach met ...] bought the house...
1 PP	**at 8 o’clock**
2 PP	*by the river* **at 8 o’clock**
3 PP	NEAR THE GYM *by the river* **at 8 o’clock**

Table 2.14: Sample Item from Jaeger et al. (2005)

One potential explanation of this finding comes from surprisal theory (Levy, 2008), according to which the parser continuously generates predictions about upcoming input based on the string of words it has already seen by allocating a certain amount of probability to all potential continuations. This in turn determines the surprisal of each potential continuation, which is inversely related to the probability of encountering the word given the preceding sentential context. The processing difficulty incurred by a given word is determined by that word's surprisal; words with higher surprisal, which are less expected, cause greater processing difficulty. In this system, antilocality effects arise because longer dependencies are associated with lower surprisal when the verb is finally encountered. Regarding Jaeger et al. (2005), this account predicts that surprisal at the matrix verb will be reduced as the number of postverbal constituents in the relative clause increases, as each additional constituent limits the number of possible continuations, and the RC must eventually come to an end. As a result of reduced surprisal, the verb is easier to process, and response times are faster at the matrix verb despite the additional distance between the matrix subject and verb.

The finding in Experiments 5 and 6 parallels Jaeger et al. (2005), because the region containing the matrix verb and subsequent material was read faster when the RC was longer. Under a surprisal view, the extra material in the RC constrains possible sentence continuations, making the matrix verb more likely and facilitating processing in the post-RC region, explaining the main effect of SIZE found in the response times. If this is an anti-locality effect, then this replication is significant beyond implicit prosody by providing more evidence for antilocality effects in English. Such effects are more frequently documented in head-final languages such as Hindi (Vasishth & Lewis, 2006), German (Konieczny & Döring, 2003; Levy & Keller, 2013), and Japanese (Nakatani & Gibson, 2008); moreover, the one study I know of that has documented anti-locality effects in English was an SPR study that was not published (Jaeger et al., 2005), meaning that convergent evidence from other sources would grant more confidence in this result.

As an alternative to anti-locality, or as a potential mechanism behind anti-locality effects, the faster response times in the Long condition could also reflect the assignment of implicit prosody. When the RC is long, readers are more likely to have already placed a prosodic boundary at the end of the RC than when this RC is short. This implicit prosody could lead to faster response times in the following region in at least two ways. First, since participants (are more likely to) have closed the prosodic phrase containing the relative clause, they could already know that they need to prepare a separate prosodic phrase for the upcoming material. This early projection and/or expectation of an upcoming prosodic phrase for the post-RC material could facilitate processing, resulting in the shorter response times in this region. An alternative but not mutually exclusive possibility is that the (higher probability of a) boundary after the long RC is a cue to the reader that the upcoming material will likely be part of a different syntactic constituent; moreover, since prosodic boundaries are known to encourage long-distance attachment, this boundary could also serve to alert the parser to the likely possibility of a long distance dependency. In this case, the parser would be more prepared for the matrix clause, again resulting in faster response times in this region.

It is important to keep in mind that I did not make any specific predictions about response times at the post-RC region, and the experiment was not designed to address anti-locality effects, so the preceding discussion is purely speculative. However, it is a useful demonstration of several points made in Chapter 1: how implicit prosody (i) is intimately tied up in classic sentence processing effects, yet (ii) has not always been considered as a potential source of these effects, with some major exceptions, e.g., RC length and attachment, and (iii) may be extremely difficult to pinpoint, due to open questions about which type of processing (e.g., syntactic, semantic, prosodic) and which stage of processing (e.g., first pass, reanalysis) are indexed by RTs at any given point. However, these questions should not be avoided just because they are difficult, and I hope to have shown that the Maze and SPR are viable tasks to investigate the online impact of implicit prosody on sentence processing.

2.11 General Discussion

In this chapter, I have established that three effects that are sensitive to implicit prosody replicate in the Maze. On the basis of Experiments 1 and 2 on homographs, I concluded that both the Maze and SPR are sensitive to metrical structure. On the basis of the NP/Z ambiguity in Experiments 3 and 4, I concluded that the Maze does not induce a boundary-heavy prosody that would counteract Late Closure. Across these four experiments, I also found that effects on response times in the Maze are larger and more localized than in SPR, replicating previous work (Boyce et al., 2020; Witzel et al., 2012, i.a.). Finally, I replicated the well-known effect of RC length on attachment preferences in both the Maze and SPR, suggesting that the implicit prosody assigned by readers in these tasks is sufficiently similar to the implicit prosody used in whole-sentence presentation, even if task demands could plausibly influence prosodic assignment.

As stated earlier, eye-tracking while reading would be the ideal method in order to study potential differences between first pass and default implicit prosody. However, eye-tracking studies were not feasible for this dissertation because of the COVID-19 pandemic. This methods chapter was necessary to establish that implicit prosodic effects can be replicated in both the Maze and SPR. Because these tasks produce comparable results to previous work and approximate the incremental nature of online listening, I will use these tasks in the rest of the experiments in this dissertation. Future work should attempt to replicate the results of the studies in Chapter 4 in eye-tracking while reading, once this is possible.

Chapter 3

The incremental assignment of implicit prosody

In Chapter 2, I established that the Maze is a viable task for studying implicit prosody. In this chapter, I return to the question of incremental prosodic assignment and the potential differences between the first pass and the final prosody. I walk through a model of how the parser would plausibly assign structure in two well-known constructions, the NP/Z garden path and the complex NP + relative clause sequence. I also provide an overview of grammatical constraints on the syntax-prosody interface. These grammatical constraints are typically used to model the prosody of complete sentence structures, i.e., the final "default" prosody of a sentence, which is distinct from the task of an incremental parser that assigns structure one word at a time, without knowing in advance the final content of the sentence. I therefore explain how the parser could incorporate these principles in order to structure incoming material word-by-word. This discussion provides the necessary background for Chapter 4, in which I extend previous work on the complex NP + RC structure by manipulating the length of the complex NP in order to address the question of incremental prosodic assignment.

3.1 Possible prosodic structures

In the rest of this dissertation, I will assume that the prosodic hierarchy consists of three categories above the foot: the prosodic word ω, the phonological phrase ϕ, and the intonational phrase ι (Ito & Mester, 2013; Selkirk, 2011). These categories (roughly) correspond to syntactic units: ω correspond to (lexical) syntactic words (X^0), ϕ correspond to syntactic phrases (XP), and ι correspond to clauses (CP). It is worth noting that previous psycholinguistic work has frequently adopted the categories of the Intermediate Phrase ($\approx \phi$) and the Intonational Phrase ($\approx \iota$) (Carlson, Clifton, & Frazier, 2009; Carlson & Tyler, 2018; Frazier, Carlson, & Clifton Jr, 2006; Lee & Garnsey, 2012, i.a.); nothing crucial rests on the distinction between these terms as long as we recognize two prosodic categories above the word, as this is primarily a difference in labels. In the following examples of how a prosodic parser would incrementally assign structure, I will assume that the parser only builds non-recursive prosodic structures in which prosodic constituents cannot dominate constituents of the same category, e.g., a ϕ cannot dominate another ϕ. This is a choice of convenience, following previous work in psycholinguistics; for a discussion of possible extensions to recursive structure, see Section 3.4.

In the following discussion, I introduce grammatical constraints that govern the syntax-prosody interface and discuss how they might be implemented by the parser. It should be emphasized that these constraints are violable, and the parser may consider parses that do not satisfy (some of) the constraints. This connects to a larger point: there is a lot of variation in prosodic structure, and there are often many licit ways to phrase any given string. Beyond the grammatical constraints discussed below, phrasing can depend on factors like speech rate and information structure. Because there are typically many possible phrasings, it is more accurate to consider the question of which phrasing the parser will pursue and what the default prosody is for any structure as a question of which phrasings are most likely.

On a related note, I describe the parser below as if it were a serial parser, pursuing a single analysis at a time. This choice is based on ease of exposition. It is also plausible that the prosodic parser builds multiple structures in parallel and assigns different probabilities to the various possible phrasings, as in the surprisal theory of syntactic processing (Levy, 2008). The reader can translate between the serial parser described below and a parallel parser as follows. When I describe the parser as building a particular parse, this can be considered the parse that is assigned the most probability by a parallel parser. When I describe the parser as engaging in reanalysis from one structure to another, for a parallel parser this would involve the reallocation of probability from the initially favored parse to a different parse.

3.2 Constraints on syntax-prosody correspondence

3.2.1 Alignment Theory and Match Theory

A common assumption in theoretical work on the syntax-prosody interface is that prosodic structure mirrors syntactic structure, but does so imperfectly; cf. direct reference theories, which do not posit a separate prosodic level of representation (Kaisse, 1985; Pak, 2008; Samuels, 2009; Scheer, 2012; Seidl, 2013). Two prominent theories in this area are Alignment Theory (Selkirk, 1986, 1996) and Match Theory (Ito & Mester, 2013; Selkirk, 2011, i.a.). Both of these theories adopt an Optimality Theoretic framework in which syntax-prosody mapping constraints enforce correspondence between syntactic and prosodic structures and are ranked with respect to wellformedness constraints governing prosodic properties like size and balance. When the wellformedness constraints are ranked above the mapping constraints, they can cause the prosodic structure to diverge from the syntax.

To my knowledge, psycholinguists tend to invoke Alignment constraints rather than Match constraints when they appeal to grammatical theories of syntax-prosody correspondence (Fodor, 2002a, 2002b; Hirose, 2003; Hwang & Schafer, 2009; Wagner & Watson, 2010; Webman-Shafran & Fodor, 2016). This trend may simply be due to the fact that Alignment constraints have been around much longer, and are therefore better known. This trend is likely also related to the use of non-recursive prosodic structures in most psycholinguistic work: Match constraints motivate recursive structure by placing stronger demands on syntax-prosody correspondence, while Alignment constraints can be satisfied by both recursive and non-recursive structures. To the extent that the parser is guided by principles resembling grammatical constraints, Alignment constraints may be easier to implement in a model of incremental processing than Match constraints, because the latter require knowledge of complete syntactic phrases in order to be evaluated, whereas the former only require the parser to know the location of constituent edges. The subtle differences between these two views of syntax-prosody correspondence will be important to consider in future theorizing on the incremental assignment of prosodic structure.

Alignment Theory and Match Theory differ primarily in the nature of the mapping constraints enforcing syntax-prosody correspondence. In Alignment theory, a family of Align constraints cares about aligning the edges of syntactic constituents to the edges of prosodic constituents. Two such constraints, ALIGN-R(XP, ϕ) and ALIGN-L(XP, ϕ), are defined in (34). Importantly, each of these constraints refers to either the left edge or the right edge, but not both. In contrast, Match constraints require syntactic and prosodic constituents to "match" - here, I adopt Elfner's (2012, 2015) terminal-based definition of matching, according to which an XP and a ϕ match if ϕ dominates all and only the phonological exponents of the terminal nodes of XP. MATCH(XP, ϕ), defined in (35), requires each XP in the syntactic structure to have a matching ϕ in the prosodic structure. Comparable constraints exist at the word level (relating ω to X^0) and at the clause level (relating ι to clauses).

(34) a. ALIGN-R(XP, ϕ): Assign one violation for every XP in the syntactic representation whose right edge is not aligned with the right edge of a ϕ in the prosodic representation.

b. ALIGN-L(XP, ϕ): Assign one violation for every XP in the syntactic representation whose left edge is not aligned with the left edge of a ϕ in the prosodic representation.

(35) MATCH(XP, ϕ): Assign one violation for every XP in the syntactic representation that is not matched by a corresponding ϕ in the prosodic representation.

To understand the differences between Alignment and Match constraints, consider the two-word right-branching syntactic structure in (36) and two possible prosodic phrasings, shown in (37a) and (37b). The prosodic structure in (37a) perfectly matches the syntactic structure in (36). It satisfies ALIGN-R(XP, ϕ) because both the right edge of XP and the right edge of YP are aligned with the right edge of at least one ϕ (in this case, both ϕ_1 and ϕ_2). Similarly, ALIGN-L(XP, ϕ) is satisfied because the left edge of XP is aligned with the left edge of ϕ_1, and the left edge of YP is aligned with the left edge of ϕ_2. Finally, MATCH(XP, ϕ) is satisfied because XP is matched by ϕ_1, as both dominate the set of terminal nodes $\{x, y\}$, and YP is matched by ϕ_2, as both dominate the same terminal node, *y*.

(36) XP

x YP

y

(37) a. ϕ_1

ω ϕ_2

x ω

y

b. $\phi_1 \phi_2$

ω ω

x *y*

In contrast, the structure in (37b) satisfies the Alignment constraints but not MATCH(XP ϕ). ALIGN-L(XP, ϕ) is satisfied because the left edge of XP is aligned with the left edge of ϕ_1, and the left edge of YP is aligned with the left edge of ϕ_2. Similarly, ALIGN-R(XP, ϕ) is satisfied because the right edges of both XP and YP are aligned with the right edge of ϕ_2. However, this structure violates MATCH(XP, ϕ), because XP is not matched by a ϕ: there is no ϕ that dominates the terminal nodes $\{x, y\}$.

I will not discuss further the particular details of these theories; see Elfner (2018) and Bennett and Elfner (2019) for recent overviews of competing approaches to the syntax-prosody interface. For present purposes, I assume that syntax-prosody correspondence constraints such as these ensure a certain level of syntactic cohesion in the prosodic structure built by the parser.

3.2.2 Syntax-prosody correspondence and the incremental prosodification of the NP/Z garden path

As an illustration of how the parser would build structure that respects syntax-prosody correspondence, consider the following example from Bader (1998), first discussed in Chapter 1, of how the parser would assign structure upon encountering the NP/Z garden path (38). Again, this garden path involves a temporarily ambiguous NP *the little boy* that could be analyzed as the direct object of the subordinate clause or the subject of the matrix clause; the parser will initially analyze it as the direct object, in accordance with Late Closure, which will necessitate reanalysis later on. Because the ι is the relevant level of structure in this example, I do not discuss the construction of ϕ, following Bader, although the parser presumably builds

ϕ inside each ι.

(38) In order to help *the little boy* put down the package he was carrying.

In (39), at Step 1, the parser encounters the beginning of the sentence, *In order to help*. At this point, the parser will recognize that it is at the beginning of a clause, and will build a corresponding intonational phrase to contain this material. Although not discussed by Bader (1998), the parser could also recognize that this clause is a subordinate clause, and that a matrix clause will follow at some point. It may project a CP corresponding to the matrix clause, as well as a second intonational phrase corresponding to the anticipated matrix clause, as shown in Step 1.

(39) Step 1: The parser encounters “In order to help...”

First pass prosody: $_{\iota}($$_{\text{CP}}$[In order to help...]) $_{\iota}($$_{\text{CP}}$[...])

In Step 2 (40), the parser encounters the ambiguous NP, *the little boy*. Following Late Closure, the parser will attach this NP as the direct object of the verb *help*. The parser can be reasonably confident that it is at the end of the subordinate clause; since the parser will try to align clause boundaries with ι boundaries, it will close the ι after *boy*.

(40) Step 2: The parser encounters “the little boy”

First pass prosody: $_{\iota}($$_{\text{CP}}$[In order to help the little boy]) $_{\iota}($$_{\text{CP}}$[...])

At Step 3 (41), the parser encounters the matrix verb *put*. Knowing that a verb cannot start the matrix clause, the parser will have to reanalyze *the little boy* as the subject of the matrix clause. However, if the parser were to continue with its first pass prosody, this would mean that an intonational phrase boundary would be placed in the middle of the matrix clause, between the subject *the little boy* and the verb *put*. To achieve alignment of clauses and ιs, the parser will have to remove the boundary after *boy* and insert a new prosodic boundary after the verb *help*, at the end of the subordinate clause. This is a clear case where the first pass prosody and the final prosody for the structure diverge.

(41) Step 3: The parser encounters "put"

First pass prosody: * $_{\iota}$($_{\text{CP}}$[In order to help] $_{\text{CP}}$[the little boy) $_{\iota}$(put...])

Reanalyzed prosody: $_{\iota}$($_{\text{CP}}$[In order to help]) $_{\iota}$($_{\text{CP}}$[the little boy put...])

The parser will then proceed through the rest of the sentence, as in (42). As far as ι go, the parser does not encounter any new clauses, so it does not need to build additional ι; placing the rest of the sentence into the second ι achieves syntax-prosody correspondence.

(42) Step 4: The parser encounters the rest of the sentence

Final prosody: $_{\iota}$($_{\text{CP}}$[In order to help]) $_{\iota}$($_{\text{CP}}$[the little boy put down the package he was carrying.])

The preceding example has highlighted several important points. First, the incremental assignment of prosodic structure in the NP/Z garden path, following Bader (1998), results in a situation where the first pass prosody is different from the final "default" prosody for the Z structure, demonstrating the importance of considering how structure is assigned word-by-word. This example has also shown how grammatical constraints requiring syntax-prosody correspondence could be implemented by the parser: knowing that certain constituents, like CP and ι, tend to align, the parser can posit prosodic structure on the basis of its syntactic analysis.

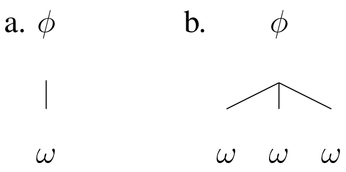
3.3 Constraints on size and balance

3.3.1 Binarity constraints

The prosodic parser is also sensitive to constraints on phrase size and balance (Fodor, 1998, 2002a, 2002b). Prosodic constituents tend to be binary; for instance, phonological phrases ϕ tend to consist of two words ω (Ghini, 1993; Selkirk, 2000, 2011). This pattern is enforced by binarity constraints. BINMIN-ϕ, defined in (43a), penalizes every ϕ that contains fewer than two ω, as in (44a). BINMAX-ϕ, defined in (43b), penalizes ϕ that contain more than two ω, as in (44b).

(43) a. BINMIN-ϕ: Assign one violation for each ϕ that dominates fewer than two ω.

b. BINMAX-ϕ: Assign one violation for each ϕ that dominates more than two ω.

(44) 

Together, these constraints conspire to achieve the ideal ϕ size of two ω. However, in the discussion that follows, I do not enforce a strict two-word limit on the structures considered by the parser. Instead, I assume that the *probability* of a boundary after a given word increases as the number of ω in a ϕ increases, provided that the last-seen word is a plausible end to the phrase, which captures the insight of much previous work showing that longer syntactic constituents are more likely to be followed by a prosodic boundary (Breen, Watson, & Gibson, 2011; Hirose, 2003; Hwang & Schafer, 2009; Watson & Gibson, 2004b, i.a.).

I do not enforce a two-word limit for several reasons. First, as explained in Section 3.1, these constraints are violable, meaning that phrasings that do not satisfy the constraints may arise. In particular, it is not uncommon for phrases to contain more than two ω, especially at faster speech rates (or in this case, presumably, faster reading rates) (Ghini, 1993; Prieto, 2005, 2006).

Additionally, there is likely to be individual variation in the maximum size of a phrase: previous work suggests that readers with a higher working memory span and who produce longer explicit prosodic phrases in speech are also more likely to have longer implicit prosodic phrases in silent reading (Bishop, 2020; Swets, Desmet, Hambrick, & Ferreira, 2007). Again, this suggests that implementing a strict two-word limit would not accurately reflect the behavior of the parser.

For these reasons, I allow for phrases containing more than two words in the following demonstrations of how the parser would assign structure incrementally. To reiterate, I still capture the spirit of the BINMAX constraint - that there tends to be a limit to how many words can be in a phonological phrase - by assuming that the probability of placing a boundary after a word increases as the number of words in a phrase increases.

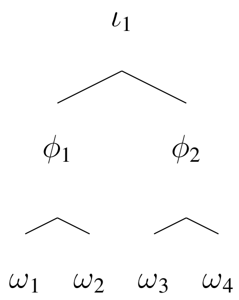
3.3.2 Uniformity and balance

Additional size constraints regulate prosodic balance: the tendency for prosodic sisters to be of the same (or similar) size. This principle is often referred to as Uniformity, defined in (45) (Ghini, 1993; Sandalo & Truckenbrodt, 2002). More recently, Bellik and Van Handel (to appear) implemented Uniformity as an OT constraint, BALANCEDSISTERS, defined in (46). This constraint is satisfied by structures like (47), in which the two sisters ϕ_1 and ϕ_2 each have the same number of ω daughters. However, it is violated by the structure in (48), in which ϕ_1 has three ω daughters but its sister ϕ_2 only has one, resulting in an imbalanced structure.

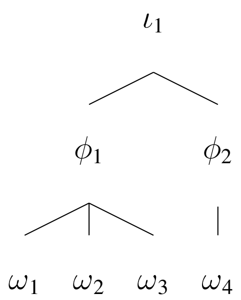
(45) Uniformity: a string is ideally parsed into same length ϕ

(46) BALANCEDSISTERS-π: Assign a violation for every node of category $\pi \in \{\iota, \phi, \omega\}$ whose children are not all of the same size. The size of a child is the number of interface nodes it branches into (ι, ϕ, ω).

(47)



(48) *



Prosodic balance is particularly relevant for present purposes, because Fodor (1998) proposed a similar constraint, the Same-Size-Sister principle, as part of her prosodic explanation of RC length effects. In her account, a long RC will search for another prosodically heavy constituent to attach to; this results in attachment to NP1, which contains both nouns and is therefore a suitably heavy attachment site to achieve prosodic balance. In contrast, when an RC is short, it will search out another prosodically light constituent, and therefore attaches to the second NP, which consists of a single noun.

3.3.3 Size, balance, and the incremental prosodification of relative clauses

As described in Chapters 1 and 2, implicit prosody has been argued to influence the attachment of relative clauses (RCs) in sentences like (49): readers are more likely to pursue high attachment of the RC to the first noun phrase "the brother" when they insert an implicit prosodic boundary after the second noun phrase "the psychiatrist." Previous studies have established the role of implicit prosody by manipulating the length of the RC: a short RC, as in (49a), is less likely to be preceded by a prosodic break than a long RC, as in (49b), and readers provide more high attachment interpretations when the RC is long (Fodor, 1998, 2002a, 2002b, et seq.).

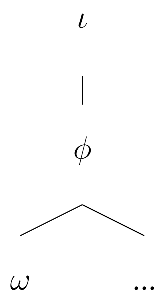
(49) a. The brother of the psychiatrist [who listened] paired a Chilean sea bass with an aggressive zinfandel.

b. The brother of the psychiatrist [who listened to the callers] paired a Chilean sea bass with an aggressive zinfandel.

I will now walk through the incremental assignment of prosodic structure to these complex NP + RC sequences in order to demonstrate how the parser's decisions could be guided by size and balance considerations, resulting in the absence of a pre-RC prosodic boundary in (49a) but not in (49b). First, consider the sentence with a short RC, (49a). In the first step of processing this sentence, (50), the parser encounters *The brother*. Since this is the beginning of the sentence, the parser will create both an ι and a ϕ in which to place *The brother*. It also leaves this ϕ open, so that more ω can be placed into this phrase; closing the phrase at this position would be dispreferred according to BINMIN, because it would result in a ϕ containing a single word. Note that I am treating all lexical words as prosodic words, and am ignoring function words, which are light, when counting the number of ω in each phrase.

(50) Step 1: The parser encounters "The brother"

First pass prosody:

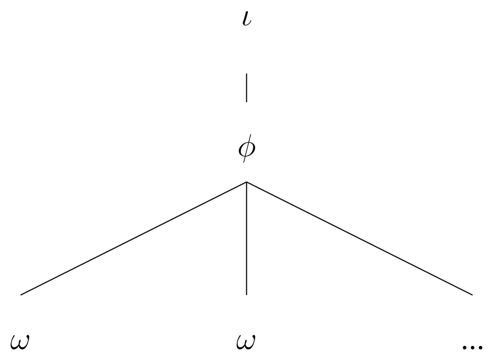


(The brother

In (51), the parser reads *of the psychiatrist*, which it places into the ϕ currently being processed. At this point, the parser could either close the ϕ, because it now contains two words, or leave it open. Drawing a parallel to syntactic parsing, I assume that the parser follows a principle I will call Prosodic Late Closure, such that its baseline preference is to leave ϕ open and attach incoming material into the ϕ currently being processed; however, because this principle is sensitive to prosodic pressures, the strength of this preference decreases as the length of the ϕ increases. The principle can also be completely overridden in the face of unambiguous evidence that the parser is at a phrase edge and needs to close the ϕ, such as a comma or a period.

(51) Step 2: The parser encounters “of the psychiatrist”

First pass prosody:



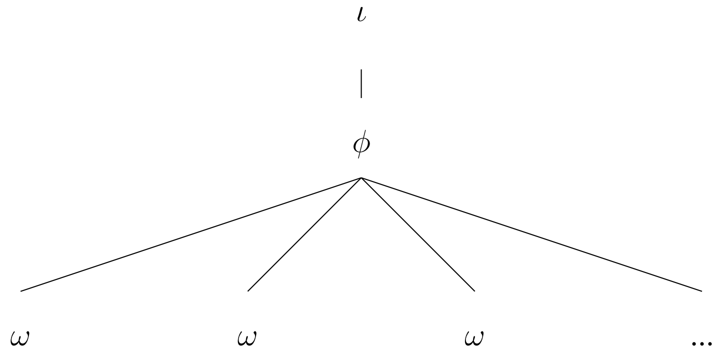
(The brother of the psychiatrist

Prosodic Late Closure allows the parser to avoid the potential downsides to closing a constituent off at a relatively early juncture. First, the parser cannot be certain at this early stage in the sentence that it has seen the entire NP: if the parser places a boundary at this position, but the NP continues, then this boundary would disrupt prosodic cohesion of the NP, which would negatively impact syntax-prosody correspondence. Second, the parser cannot be certain how long the upcoming material will be, which could motivate leaving boundaries open for several reasons. If the parser closes a boundary, and the upcoming material is a single word in a position that highly encourages or requires a prosodic boundary (e.g., at the end of the sentence, at a comma or semicolon, or at some major syntactic boundary), then the parser risks having to parse this next word into a single ϕ, violating BINMIN. Additionally, prosodic balance is presumably important to the parser, but its implementation requires at least some knowledge of how short or long several consecutive constituents are. At any moment before seeing the entire sentence, but especially at the beginning, the parser has limited access to this information. As such, in the absence of other reasons to posit a boundary (such as a strong cue to a syntactic boundary, or an overly long ϕ containing many words), the parser's default preference may be to leave a phrase open, and to revise the structure by dividing the phrase into smaller chunks later on if needed.

Next, the parser encounters the relative pronoun and the relative clause verb, *who listened*, in (52). The parser will attach this ω into the currently open ϕ. Since the parser cannot be certain that the relative clause is complete at this point, which could otherwise motivate closing the ϕ so that its right edge would align with the right edge of the relative clause, the parser will leave this ϕ open.

(52) Step 3: The parser encounters "who listened"

First pass prosody:



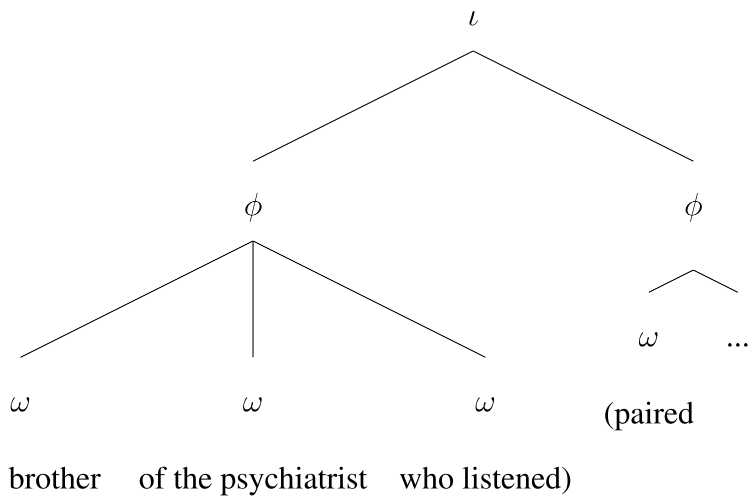
(The brother   of the psychiatrist   who listened

At Step 4, in (53), the parser reads the matrix clause verb *paired*. This verb is an unambiguous cue that the relative clause is complete. In response, the parser will close off the ϕ at the end of the relative clause, i.e., after *listened*; this achieves syntax-prosody alignment by placing the right edge of a ϕ boundary at the right edge of the RC. By putting the entire complex NP and the RC into a single phrase, rather than splitting these constituents up into two separate phrases, e.g., $_{\phi}$*(The brother of the psychiatrist)* $_{\phi}$*(who listened)*, the parser also avoids creating a ϕ that would violate BINMIN due to containing only one prosodic word. Thus, in the case of the short RC, the final prosody[1] of the complex NP + RC sequence resembles that of the first pass prosody, as both the first pass prosody at each step and the final prosody do not have a boundary between the NP and the RC.

(53) Step 4: The parser encounters "paired"

[1]Because my main goal with this example is to show how to capture the classic RC length effect with an incremental parser, I set aside what happens when the parser assigns structure to the rest of the sentence, focusing only on the complex NP and the RC. However, the rest of the matrix clause could plausibly influence the prosody of the first half of the sentence if its length requires further revisions to ensure a balanced parse.

Final prosody:

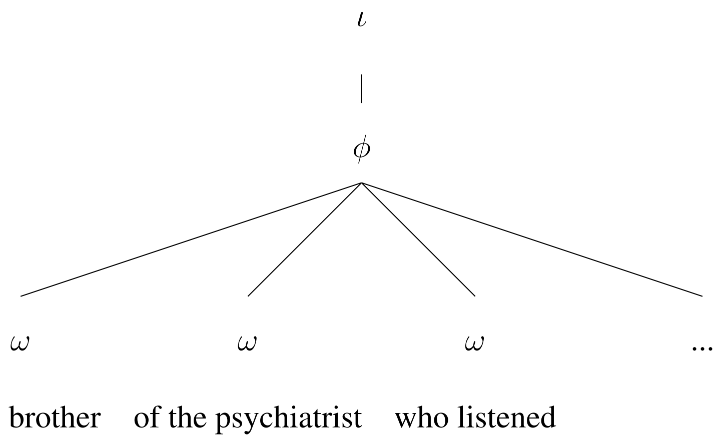


Next, consider the case of a sentence with a long RC, repeated in (54). The first three steps of parsing this structure would be identical to the short RC case, because the material is identical through the relative clause verb in both sentences. This step is repeated in (55) in order to demonstrate the parse state before the additional material in the RC is encountered.

(54) The brother of the psychiatrist [who listened to the callers] paired a Chilean sea bass with an aggressive zinfandel.

(55) Step 3: The parser encounters “who listened”

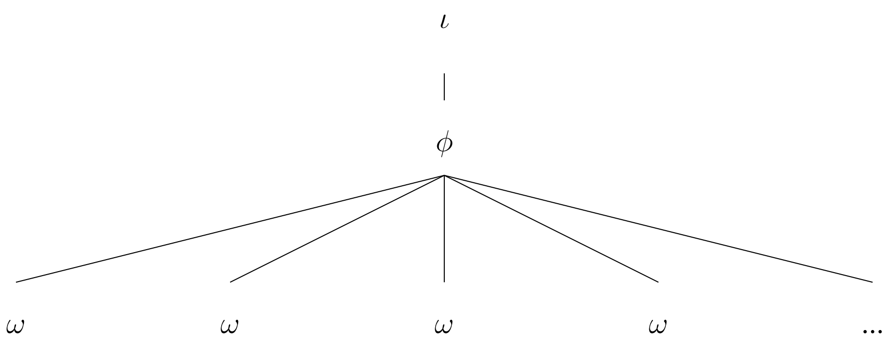
First pass prosody:



Next, the parser processes the prepositional phrase *to the callers* in (56). I assume that the parser will attach this into the current ϕ and leave this ϕ open, waiting until it knows how long the RC is before it closes off the ϕ and decides whether and how to split the material up into additional phrases in order to satisfy both length and balance pressures. Another likely structure is one in which the parser decides to close the ϕ at this position, after the PP; however, following Prosodic Late Closure, I contend that there would still be a preference to leave the ϕ open, because there could still be additional incoming material in the RC that should be prosodically phrased with the rest of the RC, and the parser may decide it is better to wait to split up the ϕ once it is certain of the RC's length.

(56) Step 4: The parser encounters “to the callers”

First pass prosody:

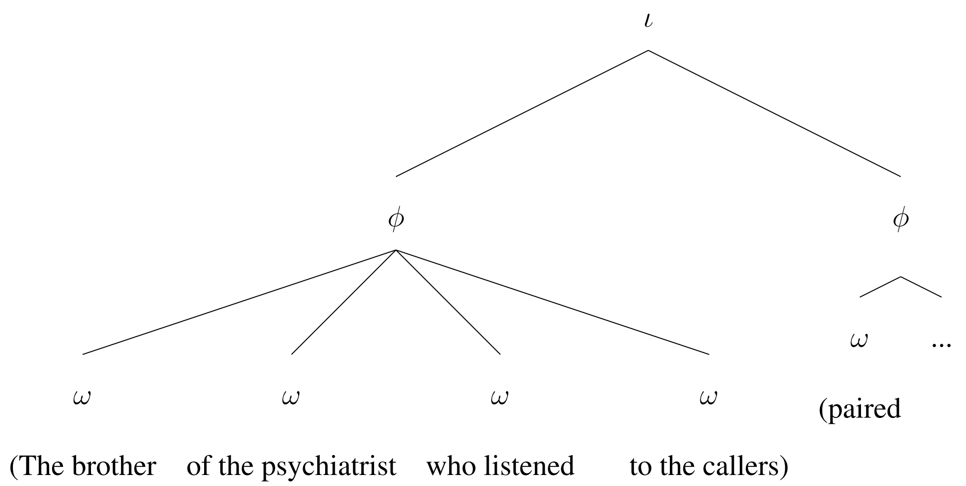


(The brother of the psychiatrist who listened to the callers

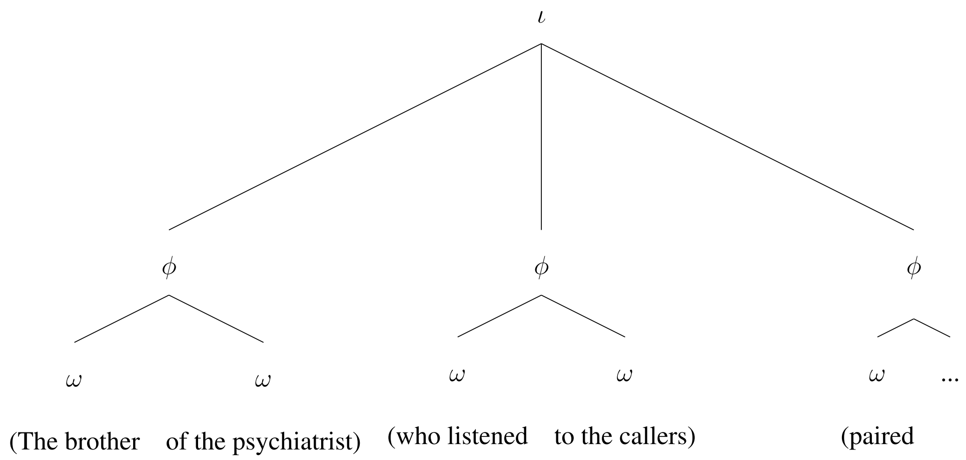
At the next step in (57), the parser encounters the matrix clause verb *paired*, which again is an unambiguous cue that the relative clause is complete. The parser will now close the ϕ between the PP and the verb. The parser may decide to leave this ϕ as is, as in (57a); although this ϕ violates BINMAX, because it contains more than two words, leaving the ϕ also avoids an additional revision process, which is potentially costly. Alternatively, the parser could decide that it would like to split up the ϕ into fewer phrases in order to avoid the BINMAX violation. In this case, the parser would likely place a boundary between the complex NP and the RC, as in (57b): this would result in two ϕ that each contain two words, satisfying the binarity constraints. These two ϕ would also satisfy BALANCEDSISTERS, because the sister phrases each contain two ω.

(57) Step 5: The parser encounters "paired"

a. Final prosody, option 1:



b. Final prosody, option 2:



I assume that both (57a) and (57b) are possible structures that the parser could build; one way to achieve optionality is to stipulate that the parser stochastically pursues one of the two analyses, as in the Unrestricted Race Model of syntactic parsing (Van Gompel, Pickering, Pearson, & Liversedge, 2005; Van Gompel, Pickering, & Traxler, 2000). However, it is important to point out the relationship between first pass prosody and the final prosody in each case. When the parser pursues the first option, (57a), in which the NP and the RC are in a single phrase, then the first pass prosody, which never placed a boundary between the NP and the RC, would not differ from the final prosody. When the parser pursues the second option, (57b), then the first pass prosody and the final prosody would differ, because only the latter has a ϕ boundary between *of the psychiatrist* and *who listened*. Again, this illustrates a major point from Chapter 1: the first pass prosody may not always match the final "default" prosody associated with a structure, and only by considering the incremental assignment of prosodic structure can we determine when they will diverge.

These potential differences between the first pass and the final prosody also have implications for the parser's attachment decisions. Under the assumption that the first pass prosody never places a boundary after the NP and before the RC, then at the time that the parser encounters the relative pronoun and first makes an attachment decision, as in Step 3 (55), the prosody would favor low attachment following Fodor's account, because there would be no pre-RC boundary. When the final prosody also lacks a boundary, as in (57a), then the final prosody would also favor low attachment due to the absence of a boundary. However, when the final prosody changes and includes a pre-RC boundary, as in (57b), then the prosody would favor high attachment, and the prosody could encourage the parser to revise its initial attachment decision (or to re-allocate probability to the high attachment parse, for a parallel parser).

Finally, compare the predicted final prosodic structures for the short RC and the long RC sentences, repeated in (58). For the short RC, the most likely parse is (58a), in which no boundary separates the complex NP and the RC. For the long RC, there are two likely parses: one in which no boundary separates the complex NP and the RC, (58b), and one in which there is a pre-RC boundary, (58c). Thus, this toy model of an incremental parser captures Fodor's account of the RC length effect: a boundary between the NP and the RC is more common when the RC is long. Following her account, this pre-RC boundary should encourage non-local attachment, such that the more frequent presence of a pre-RC boundary in the long RC condition will result in more high attachment interpretation responses.

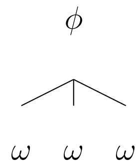
(58) a. $_{\phi}$(The brother$_{\omega}$ of the psychiatrist$_{\omega}$ who listened$_{\omega}$)

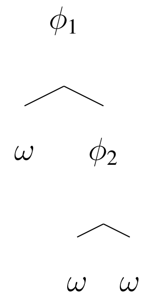
b. $_{\phi}$(The brother$_{\omega}$ of the psychiatrist$_{\omega}$ who listened$_{\omega}$ to the callers)

c. $_{\phi}$(The brother$_{\omega}$ of the psychiatrist$_{\omega}$) $_{\phi}$(who listened$_{\omega}$ to the callers$_{\omega}$)

3.4 Extensions to recursive structure

In this chapter, I have only considered prosodic parses that are non-recursive, i.e., that do not have any instances of a prosodic constituent dominating another constituent of the same category. Such a structure is exemplified in (59a): the ϕ does not dominate any other ϕ, and is therefore non-recursive. In contrast, an example of a recursive structure is shown in (59b), in which the outermost phonological phrase, ϕ_1, dominates another phonological phrase, ϕ_2.

(59) a. 

b. 

The existence of recursive prosodic structure is a matter of debate in phonological theory. Some researchers have argued for recursion at the ω level (Bennett, 2018; Booij, 1996; Ito & Mester, 2009; Peperkamp, 1997), the ϕ level (Elfner, 2015; Elordieta, 2015; Ito & Mester, 2012, 2013), and the ι level (Féry, 2010; Ladd, 1986; Myrberg, 2013). Others contend that prosodic structure is non-recursive, and that this is a crucial distinction between syntax and prosody (Jackendoff & Pinker, 2005; Nespor & Vogel, 1986; Selkirk, 1984; Vogel, 2009). The choice to only consider non-recursive structures in this dissertation should **not** be interpreted as a rejection of prosodic recursion. Rather, this choice is a matter of convenience: as described by Bellik and Kalivoda (2019), the number of possible prosodic parses explodes when recursive structure is allowed. For instance, a string of 4 words has 8 possible prosodic parses under Strict Layering (i.e., with only non-recursive parses allowed), but hundreds of possible parses when recursive structures are permitted. Thus, excluding recursive parses vastly simplifies the set of possible phrasings, which facilitates the generation of predictions about the implicit prosodic structure assigned by the parser.

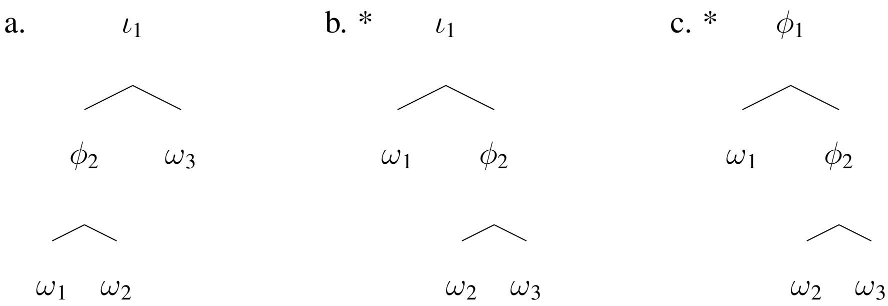
Moreover, prosodic recursion is typically motivated by close examination of domain-specific processes and gradient phonetic phenomena that require reference to different levels of embedding of a given prosodic category. However, it is unclear whether these kinds of phonetic processes are represented implicitly during silent reading. Moreover, our current understanding of implicit prosody is such that we are still trying to figure out how to productively induce the presence or absence of *any* boundary. Without a better understanding of the relationship between implicit prosodic structure and behavioral measures like response times, our psycholinguistic methods are likely not sensitive enough to distinguish different levels of recursive structure. More generally, it is often difficult to distinguish non-recursive and recursive parses, because they can sometimes achieve the same effect: for instance, both the non-recursive (59a) and the recursive (59b) place N2 at the right edge of a ϕ, and so evidence for a prosodic boundary after N2 would be compatible with either structure. Thus, given the overwhelming number of parses made possible by the introduction of recursion, as well as the potential difficulty of distinguishing recursive and non-recursive parses at this stage of our understanding, only non-recursive parses are considered here.

The choice to only consider non-recursive structures follows previous psycholinguistic work: most (if not all) research on prosodic processing has assumed non-recursive prosodic structure. However, as the theoretical debate over the existence of prosodic recursion advances, any theory of (implicit) prosodic processing must eventually decide whether and how to incorporate recursive prosodic structure.

Several additional constraints from the phonological literature could become relevant to theorizing about the prosodic parser if and when recursive structures are entertained. One such constraint, STRONGSTART, is defined in (60) (Elfner, 2012; Selkirk, 2011). STRONGSTART requires the first daughter of each prosodic constituent to be of the same level or higher than the sister constituent to its immediate right. This constraint is satisfied by a structure like (61a), in which the first daughter of ι_1, ϕ_1, is of a higher category than the sister to its right, ω_3. However, it is violated by structures like (61b), in which the first daughter of ι_1 is ω_1, which is lower on the prosodic hierarchy than its sister, ϕ_2. The recursive structure in (61c) also violates STRONGSTART, for a similar reason: the first daughter of ϕ_1 is ω_1, which is weaker than its sister ϕ_2.

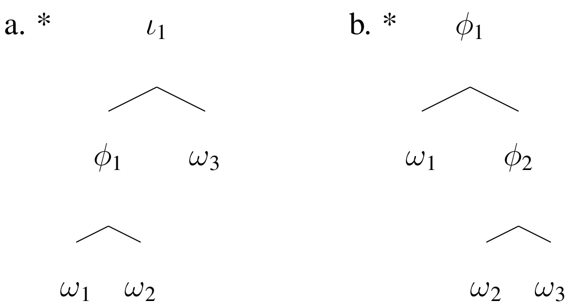
(60) STRONGSTART: Assign one violation for every prosodic constituent whose leftmost daughter constituent is lower in the Prosodic Hierarchy than its

sister immediately to its right.

(61)


Another constraint on sister constituents, EQUALSISTERS, is defined in (62)[2] (Bellik, to appear; Ito & Mester, 2019; Myrberg, 2013). This constraint penalizes any structure in which sister nodes belong to different categories, such as (63a), in which ϕ_1 is the sister of ω_3, and (63b), in which ω_1 is the sister of ϕ_2.

(62) EQUALSISTERS: Assign one violation for each pair of sister nodes that are not of the same prosodic category.

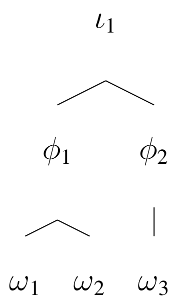
(63)


Note that STRONGSTART differs from EQUALSISTERS, in that the former would not penalize (63b): STRONGSTART is okay with differences in category as long as the leftmost sister is of a higher category than the sister to its right. Finally, there is an important distinction between EQUALSISTERS and BALANCEDSISTERS: EQUALSISTERS

[2]Other implementations of EQUALSISTERS are possible; see Bellik (to appear) for discussion.

would be satisfied by (64), because the two sisters ϕ_1 and ϕ_2 belong to the same category, but BALANCEDSISTERS would penalize this structure, because the two sisters have a different number of children.

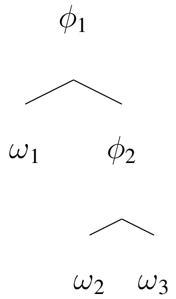
(64)



In this chapter, STRONGSTART and EQUALSISTERS were not relevant, because they only assign violations in structures where sisters belong to different prosodic categories, whereas I have only considered non-recursive structures and structures without level skipping, in which ι only have ϕ daughters, and ϕ only have ω daughters, such that that sister nodes always belong to the same category. However, future work incorporating these more complex prosodic structures should consider how an incremental parser could be guided by these constraints. For instance, if the parser knows that weak starts are dispreferred, it might assign less probability to parses in which an ω is followed by a ϕ sister. Similarly, if the parser knows that sister nodes tend to belong to the same category, it could form predictions about upcoming structure, e.g., that a ϕ is likely to be followed by another ϕ.

The recognition of recursive structure may also require a reconsideration of the way size effects are implemented in the prosodic parser. In the preceding discussion, I have followed the definition of BINMAX according to which a ϕ must contain no more than two ω. However, analyses assuming recursive structures have proposed alternative formulations of BINMAX according to which the binarity limit is placed on the number of branches a ϕ has rather than the total number of ω that it contains (Bellik & Kalivoda, 2016; Elfner, 2012; Ito & Mester, 2019, i.a.); see Bellik and Van Handel (to appear) for a typological survey and discussion of how both types of constraints are likely needed for full empirical coverage. Under this view, a structure like (65) would not violate BINMAX: even though ϕ_1 contains three ω, it branches into only two constituents, ω_1 and ϕ_2.

(65)



Earlier in this chapter, I chose not to implement BINMAX as a strict limit on two words per ϕ, arguing instead that upper size limits should be implemented by increasing the likelihood that an implicit boundary will be inserted as the length of the ϕ currently being processed increases and the last-seen word could end the phrase. The decision to eschew a strict size limit was motivated by several considerations: grammatical constraints like BINMAX are violable, and the average size of phrases frequently depends on factors like speech rate and working memory. Although recursive structures have been set aside in this dissertation, the alternative definition of BINMAX provides additional justification for entertaining structures that exceed a strict two-word limit, because this definition would allow for ϕ containing more than two words as long as there is additional internal structure to ensure a binary-branching structure, as with ϕ_1 in (65).

3.5 General Discussion

In this chapter, I walked through how an incremental parser would structure the NP/Z garden path and the complex NP + RC sequence. The goal of this chapter was to emphasize once more the potential differences between first pass prosody and the final default prosody: determining which implicit prosody the parser will assign requires going beyond applying grammatical constraints to the final string to determine which phrasing would best satisfy constraints on syntax-prosody correspondence, size, and balance, because the parser only has access to information about a fragment of the sentence when it begins to assign structure incrementally. To address this issue, I demonstrated how the parser could use its knowledge of these grammatical constraints to structure a sentence incrementally. Finally, I have identified several areas where future work could incorporate recent advances in phonological theory, such as by considering recursive parses and the role of constraints like STRONGSTART and EQUALSISTERS that could be used to form predictions about upcoming material.

Chapter 4

Implicit boundaries and relative clause attachment

4.1 Investigating incremental assignment of prosody

In Chapter 3, I walked through several examples of the incremental assignment of prosodic structure, providing an overview of several major grammatical constraints on the syntax-prosody interface and explaining how a parser could implement these principles in online structure building. Throughout this discussion, I have stressed the importance of considering the incremental assignment of implicit prosody and the potential differences between the first pass prosody and the final prosody.

In order to investigate the timing of implicit prosodic assignment, studies must meet two desiderata: they must (i) manipulate material *prior* to the location of the hypothesized boundary, such that any effect can be detected at this location in online measures like response times, and (ii) employ methods that can track incremental processing. To that end, this chapter presents a series of experiments investigating the effects of constituent length on the assignment of implicit boundaries by manipulating the length of the complex NP via modification of N1 and N2 in the Maze.

By manipulating the length of the complex NP, the experimental design ensures that any effect of length on the presence of a boundary occurs before the location where that boundary would be assigned. By using the Maze, the experiments can track incremental processing and determine whether and how length-induced boundaries affect online parsing. The experiments in this chapter should be seen as a first step toward answering questions about timing. Such questions are complex, and developing a complete theory of implicit prosodic assignment is outside the scope of this dissertation. This study should serve as a guide for future research in this area by highlighting potential challenges, open questions, and experimental design considerations.

The present experiments also refine our understanding of how implicit prosody and information structure interact. I return to the Repellent Boundaries Hypothesis, the Focus Attraction Hypothesis, and the Visibility First Hypothesis from Chapter 1. Across the three experiments, I show that neither boundaries nor focus alone can explain the effects of modification on attachment responses. I contend that the results are better captured by the Visiblity First Hypothesis, which states that focus only affects attachment preferences when two attachment sites cannot be distinguished on the basis of prosodic visibility.

4.2 Inducing boundaries with NP modification

In a series of experiments, I extend our understanding of length effects in silent reading by manipulating the length of the complex NP. Length was manipulated by modifying either of the two nouns in the complex NP; compare (66a), in which neither noun is modified, to (66b), in which N1 is modified, and to (66c), in which N2 is modified. Anticipating the order of the experiments, I focus on the contrast between no modification and N2 modification (66b) in this section; I return to N1 modification in the presentation of Experiment 8.

(66) a. The niece of the waitress [who chided herself over the blunder] had just turned seventeen.

b. The **incredibly diligent** niece of the waitress [who chided herself over the blunder] had just turned seventeen.

c. The niece of the **incredibly diligent** waitress [who chided herself over the blunder] had just turned seventeen.

It is well known that longer constituents are more likely to be followed by a prosodic break, both in overt and implicit prosody (Breen et al., 2011; Hirose, 2003; Hwang & Schafer, 2009; Watson & Gibson, 2004b). Thus, I hypothesized that introducing a two-word adjective phrase (*incredibly diligent*) before N2 (*waitress*), as in (66b), would make readers more likely to place an implicit prosodic boundary between the second noun and the relative clause. This fulfills the first desideratum described in the previous section: the experiment introduces a manipulation that occurs *before* the location of the hypothesized pre-RC boundary, such that one could potentially track the impact of this boundary on first pass parsing.

In order to achieve the second desideratum - an experimental design that allows us to track incremental parsing - I ran the N2 modification study in the Maze (Experiment 7). I also chose a well-known sentence processing effect that can track whether participants preferentially pursued a high attachment or low attachment parse: the ambiguity advantage. In (67), the reflexive pronoun determines whether the RC must attach high or low. In (67a), the pronoun *herself* matches the gender of both nouns in the complex NP, *niece* and *waitress*. However, in (67b), the reflexive *herself* requires a low attachment parse of the RC, because it is only compatible with N2, *waitress*, but not N1, *son*. In contrast, in (67) the reflexive *himself* requires a high attachment parse because it is only compatible with N1, *son*. Previous work has demonstrated an ambiguity advantage, such that reading times at the reflexive are fastest in the Ambiguous condition (67a) and slowest in the High Attachment condition (67c) (Sloggett, Van Handel, Sasaki, et al., 2020; Swets et al., 2008; Traxler, Pickering, & Clifton Jr, 1998; Van Gompel et al., 2005, i.a.).

(67) a. The **niece** of the **waitress** [who chided **herself** over the blunder] had just turned seventeen.

b. The son of the **waitress** [who chided **herself** over the blunder] had just turned seventeen.

c. The **son** of the waitress [who chided **himself** over the blunder] had just turned seventeen.

Various proposals have been put forth to explain the source of the ambiguity advantage; in this chapter, I remain agnostic as to the source of this effect. For present purposes, what matters is that high attachment is found to be most costly, as reflected in longer response times. Thus, if I can induce an implicit prosody that favors a boundary between the complex NP and the RC, and this boundary affects attachment choices in first pass parsing, then the relative difficulty of low and high attachment when N2 is and is not modified should be reflected in differential reading times across conditions. I lay out specific predictions below; crucially, the ability to track potential influences of implicit prosody on first pass attachment choices using response time data fulfills the second desideratum.

4.3 Predicted effects of modification on implicit prosody

As a reminder, any account of implicit prosody must address (i) how a given manipulation affects the assignment of implicit prosody, and (ii) how the implicit prosody (and, potentially, other features introduced by the manipulation) affects the assignment of syntactic structure and later interpretation processes. In this section, I address the former by walking through the incremental assignment of prosodic structure for the experimental items, building on the demonstrations from Chapter 3. I then revisit the Repellent Boundary Hypothesis, the Focus Attraction Hypothesis, and the Visibility First Hypothesis from Chapter 1 in Section 4.4 in order to address their predictions about how the proposed prosodic structures would influence attachment.

The complex NP + RC sequence from (67) is repeated in (68); I use curly braces to indicate the optionality of the length manipulation in order to distinguish this from the parentheses used to indicate prosodic phrasing. As described above, I hypothesize that increasing the size of a constituent increases the likelihood that it will be followed by a boundary (Breen et al., 2011; Hirose, 2003; Hwang & Schafer, 2009; Watson & Gibson, 2004b). Thus, for the Short condition of the experiment, I expect at least two phrasings to be possible: one in which the complex NP phrases separately from the RC, as in (69a), and one in which the complex NP phrases with the RC, as in (69b). However, in the Long condition, I expect that modifying N2 will increase the likelihood that a boundary is placed after the complex NP, which is now several words longer. As such, I expect that (70a), in which the NP and the RC phrase separately, will be a plausible parse, while a phrasing without a boundary between N2 and RC, as in (70b), will be much less likely in the Long condition.

(68) The niece$_{\text{N1}}$ of the {incredibly$_{\text{Adv}}$ diligent$_{\text{Adj}}$} waitress$_{\text{N2}}$ who chided$_{\text{V}}$ herself$_{\text{Refl}}$ over the blunder$_{\text{PP}}$...

(69) a. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$) % $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

b. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

(70)

a. $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$) % $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

b. ?? $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

It should be pointed out that there may still be smaller phrase breaks, but not a major boundary, between N2 and the RC in (69b). The main point is that the length manipulation makes the consistent placement of a strong prosodic boundary between N2 and the RC much more likely when the complex NP is lengthened via N2 modification in (70). For the rest of the chapter, I will continue to discuss the difference between the Short and Long condition as a difference in the likelihood of the presence or absence of this major prosodic break, but the reader should keep in mind that the length manipulation may also result in a stronger boundary in the Long condition.

While these differences in phrasing are plausible, I have emphasized throughout this dissertation that there are often multiple licit ways to phrase any given string, and length is not the only consideration that determines which words phrase together. With this in mind, I next walk through my assumptions about how an incremental parser would structure the experimental stimuli in order to demonstrate why I expect the parses in (69a,b) and (70a) to be the most likely.

4.3.1 Parsing the predicted structures

I will now consider (a simplified version of) how the parser would incrementally structure the sentences with N2 modification. As a simplifying assumption, I will stipulate that the parser only considers parses that respect a certain degree of syntax cohesion in prosodic phrasing, such that (i) the modifier phrase will always be parsed into a prosodic phrase with the noun that it modifies, e.g., (i) $_{\phi}$(*the niece*$_{\omega}$ *of the incredibly*$_{\omega}$ *intelligent*$_{\omega}$) $_{\phi}$(*waitress*$_{\omega}$) would be an unlikely phrasing, and (ii) the entirety of the relative clause is always in the same phrase, i.e., $_{\phi}$(*the niece*$_{\omega}$ *of the the waitress*$_{\omega}$ *who chided*$_{\omega}$) $_{\phi}$(*herself*$_{\omega}$ *over the blunder*$_{\omega}$) would also be unlikely. This assumption reflects the pressure for syntax-prosody correspondence represented by Align and Match constraints in Chapter 3.

One could rightly question whether the parser enforces syntax-prosody correspondence so strictly: perhaps the parser entertains phrasings that do not obey this level of syntactic cohesion, especially given that the parser's expectations of which words belong in a phrase together at any given stage of incremental processing will be influenced by a syntactic analysis that is itself incomplete and subject to revision. However, by assuming that the parser is more likely to build structures respecting syntax-prosody correspondence, the space of possible parses is circumscribed, facilitating the comparison of potential prosodic outputs and the generation of predictions for the present study.

With this assumption in place, there are three places where the parser could place a break in the complex NP + RC sequence in (71) and (72) while respecting this level of syntax-prosody correspondence: after the first noun, *niece*, after the second noun, *waitress*, and after the PP, *over the blunder*. I will now walk through the parser's decisions at the point where it has encountered each of these words. Note that for ease of exposition, I have made several other simplifications: I am eliding the work of the syntactic parser, and I am only showing what the prosodic parser does at each potential boundary location instead of showing each word-by-word step.

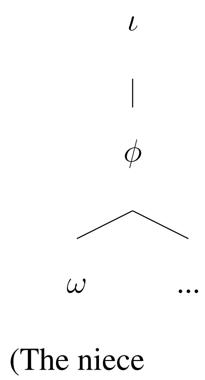
(71) The niece of the waitress who chided herself over the blunder...

(72) The niece of the incredibly diligent waitress who chided herself over the blunder...

In both the Short and Long versions of the sentence, the parser will create a new ι and a new ϕ upon encountering N1. Its first decision is whether to leave the ϕ open, as in (73a), or to close it off and place a ϕ boundary after *niece*, as in (73b). Following the principle of Prosodic Late Closure from Chapter 3, I assume that the parser will leave this ϕ open, because closing it too early would otherwise result in a ϕ containing a single word, in violation of BINMIN, which requires each ϕ to contain at least two words.

(73) Step 1: The parser encounters "The niece"

a. $_\phi$(The niece$_\omega$ ...)



b. $_{\phi}$(The niece$_{\omega}$)

ι

|

ϕ

|

ω

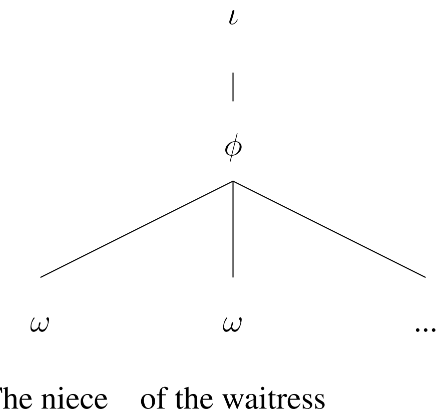
(The niece)

Next, the parser encounters the second noun phrase. In the Short condition, this consists of a single prosodic word, the noun *waitress*. Again, the parser can either leave the ϕ open, as in (74a), or close the ϕ after the second word, as in (74b)[1]. Following Prosodic Late Closure, I assume that the parser prefers the phrasing in (74a), for the same reasons that motivated this principle in Chapter 3: positing a boundary too early is risky, because the parser does not know if it has seen the entire NP or how long the upcoming material is. If the parser commits to a boundary too early, then it risks having to engage in additional revisions later on. I assume that this revision process, which would presumably involve both removing a boundary and creating a new one, would be more costly than the alternative of leaving the ϕ open and potentially returning to place a boundary later on.

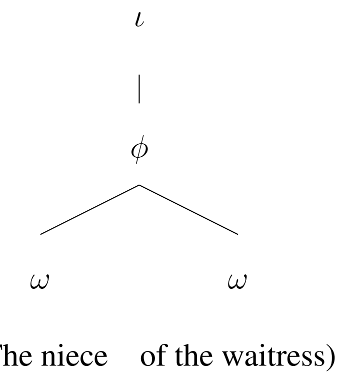
(74) Short Condition, Step 2: The parser encounters "of the waitress"

[1]The parser could also renege on its previous decision and place *the niece* in its own ϕ, starting a new ϕ with *of the waitress*, e.g., $_{\phi}$(The niece) $_{\phi}$(of the waitress ...). I find this option unlikely because it would violate the minimal size requirement by creating a ϕ containing just one prosodic word, so I do not show it here.

a. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ ...)



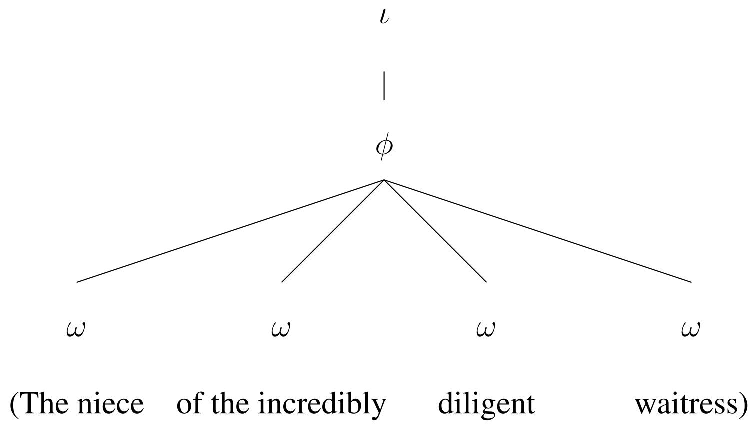
b. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$)



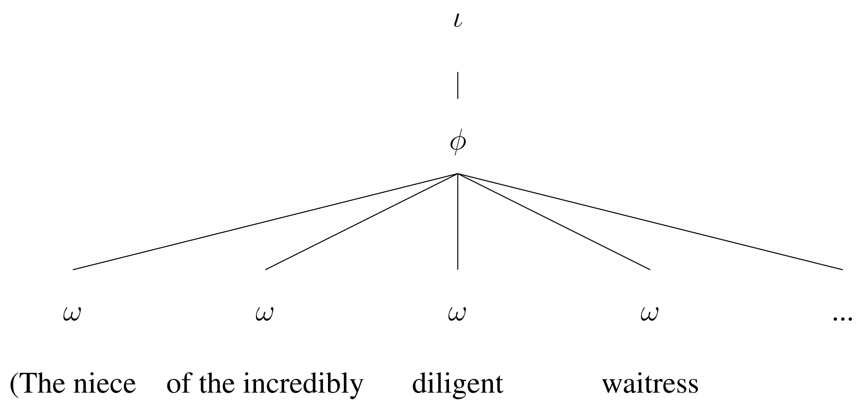
In the Long condition, the parser again has the choice to close off the ϕ after *waitress*, as in (75a), or leave this ϕ open, as in (75b). In the Long condition, this ϕ is already quite long by the time the second noun is reached, because it now contains four ω instead of just two. Again, assuming that the probability of positing a boundary is higher as phrase length increases and the last-seen word is a plausible end to the phrase, then the parser should be much more likely to place a boundary after *waitress* in the Long condition compared to in the Short condition. In other words, I expect the parser to most frequently pursue the analysis in (75a).

(75) Long Condition, Step 2: The parser encounters “of the incredibly diligent waitress”

a. $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$)



b. $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$ ...)



At this point, it is also worth returning to my reasons for assuming that the modifier, *incredibly diligent*, always phrases with the noun it modifies, N2. As noted earlier, if the parser were to split up the modifier to achieve a balanced parse, i.e.,

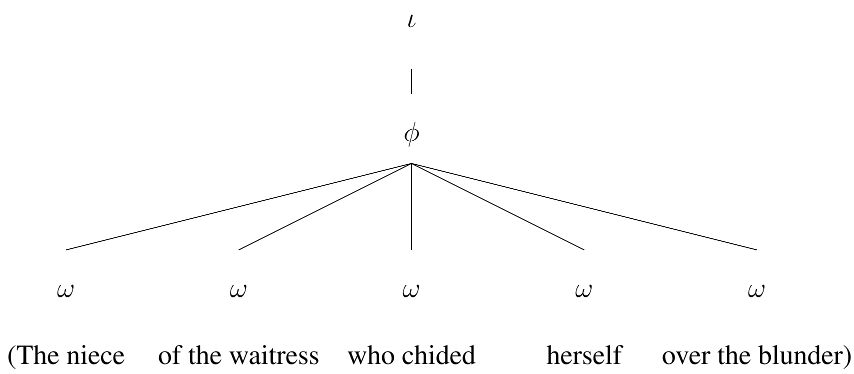
$_{\phi}$(the niece of the incredibly) $_{\phi}$(diligent waitress), this would greatly disrupt syntax-prosody correspondence. Moreover, I have the intuition that this phrasing is rather unnatural, at least compared to other options (?/* *(the niece$_{\omega}$ of the incredibly$_{\omega}$) (diligent$_{\omega}$ waitress)*).

Next, the parser encounters the RC. As a reminder, I assume that all of the elements in the RC phrase together, instead of, e.g., splitting the verb *chided* off to phrase with the complex NP. This is a safe assumption, because verbs tend to phrase with their arguments (Clemens, 2014; Selkirk, 1984). Phrasing all elements of the RC together also also avoids parses in which one of the words in the RC is phrased by itself. For instance, the phrasing $_{\phi}$(chided$_{\omega}$ herself$_{\omega}$) $_{\phi}$(over the blunder$_{\omega}$) would violate BINMIN because the ϕ corresponding to the PP would only contain a single prosodic word.

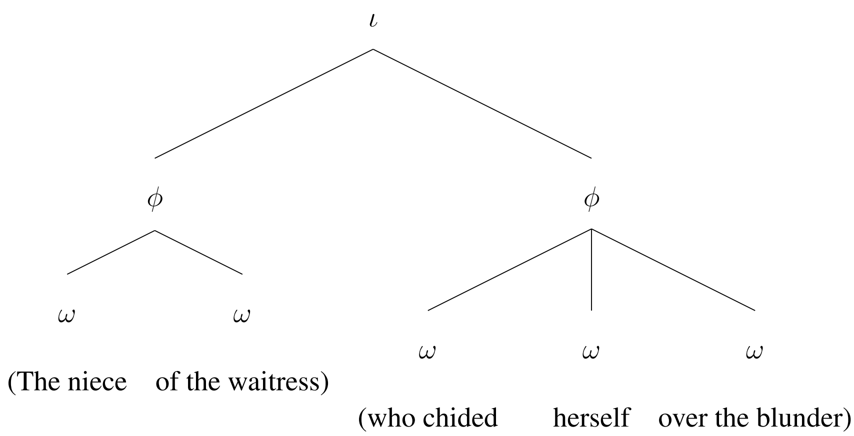
In the Short condition, the parser will start by phrasing the incoming RC with the preceding complex NP, as in (76a); because this ϕ is quite long, with five prosodic words, there is a high probability that the parser will also close off the ϕ at this point. It is also plausible that at least some of the time, the parser will decide that this ϕ is starting to become too long, and will retroactively place a boundary between the second noun, *waitress*, and the relative clause, as in (76b). This structure avoids splitting up major syntactic XPs by only placing a boundary between the complex NP and the RC. It also avoids the severe BINMAX violation of (76a) and is relatively balanced, because the sister ϕ are close in size with two and three words each.

(76) Short Condition, Step 3: The parser encounters "who chided herself over the blunder"

a. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)



b. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)



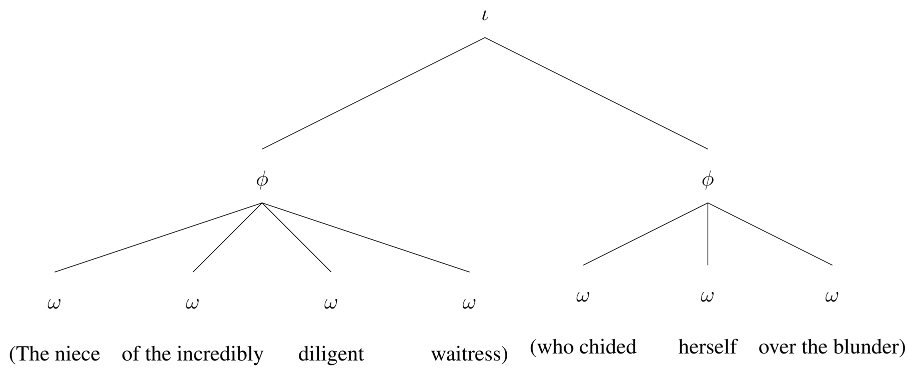
I assume that this revision from (76a) to (76b) is optional and does not always occur. As mentioned frequently throughout this dissertation, there are often many ways to phrase any given string. It is also known that readers differ in the size of the prosodic phrases they use, such that some readers will prefer shorter phrases and others longer ones (Bishop, 2020; Swets et al., 2007). Thus, I assume that prosodic phrases are allowed to contain more than two words before the parser posits a major prosodic break; though five words is a large phrase, it may be a perfectly suitable chunk for at least some readers. Moreover, there is some evidence that prosodic revision, and specifically introducing a boundary where none was previously posited, incurs a processing cost Bader (1998). It is therefore plausible that the parser would at least sometimes stick with the five-word phrase in (76a) rather than retroactively introduce a boundary between the NP and the RC, as in (76b). In summary, I assume that the parser most frequently pursues either of the two parses in (76), such that there is sometimes, but not always, a pre-RC boundary.

In the Long condition, recall that I assumed that the parser has already closed off the ϕ containing N1 and the modified N2 due to its length. The parser will continue with this analysis, building a new ϕ in which to place the RC. Thus, I expect that in the Long condition, the parser will most frequently end up with an analysis where a boundary intervenes between N2 and the RC, as in (77).

(77) Long Condition, Step 3: The parser encounters “who chided herself over the blunder”

$_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$

over the blunder$_\omega$)



I have now shown how an incremental parser sensitive to grammatical pressures governing syntax-prosody correspondence, size, and balance could plausibly assign the implicit prosodic structures that I have predicted for Experiment 7. Before discussing the potential effects of these prosodic structures on attachment and interpretation, it is important to acknowledge once more that there is so much variation in acceptable prosodic structures for a given string that it is not always the case that the final prosodies that I have proposed as the "default" will always be what gets assigned. Nevertheless, the idea of a default or most preferred final parse has long been useful in the implicit prosody literature in order to generate predictions.

In the discussion of Experiments 7 and 8, I propose possible ways to test whether the predicted prosody is indeed what readers assign when reading these sentences. Even if the proposed prosodic phrasings turn out to be selected less often than we might have expected, this discussion has laid valuable groundwork for future research by providing an overview of the constraints guiding the parser's choices and making explicit the assumptions about the relative role and timing of pressures related to syntactic cohesion, size, and balance, which could be modified in future work. Finally, this toy model has extended previous work by translating insights from constraint-based theories into an incremental parser to generate predictions about implicit prosody, rather than only discussing the default prosody in terms of what would be the best prosody for a complete string that the parser does not initially have access to. Production models of prosodic phrasing have also noted issues with models that require a certain degree of look-ahead (Breen & Clifton, 2011; Watson & Gibson, 2004b), and future work should make closer contact with this literature, which likely has additional insights for implicit prosody. However, it should also be kept in mind that there is a difference between a speaker prosodifying a sentence that they are planning versus a reader assigning structure to a sentence that they did not generate.

4.4 Predicted effects of modification and phrasing on syntax and interpretation

In Chapter 1, I presented several hypotheses about how prosodic boundaries and other information like focus could influence attachment. I review these hypotheses here before going over their specific predictions for the present study.

4.4.1 The Repellent Boundaries Hypothesis

The Repellent Boundaries Hypothesis, repeated from Chapter 1 in (78), states that the parser is less likely to target a potential attachment site when it is immediately followed by a prosodic boundary, because this boundary is taken to signal that this site and the following material are not grouped together syntactically. This hypothesis has precedent in many other theories that take boundaries to encourage non-local attachment (Fodor, 1998, 2002b; Maynell, 1999; Schafer, 1997; Watson & Gibson, 2004a, 2005).

(78) **Repellent Boundaries Hypothesis**: The parser is less likely to target a potential attachment site when it is immediately followed by a prosodic boundary, because boundaries can signal the end of a syntactic constituent, which decreases the likelihood that a word preceding a boundary would be syntactically grouped with incoming material.

4.4.2 The Modification Attraction Hypothesis

The Focus Attraction Hypothesis, restated in (79), says that the parser is more likely to target an attachment site that is focused; note that this hypothesis does not make any explicit claims about whether or how focus would interact with the presence of prosodic boundaries.

(79) **Focus Attraction Hypothesis**: It is more likely that a phrase that is neither a complement nor syntactically obligatory will be taken to modify a phrase P if P is focused than if it is not, grammatical and pragmatic constraints permitting.

The Focus Attraction Hypothesis is relevant for the present study because the modifier could cause comprehenders to interpret the modified noun as being focused by inviting them to consider a set of alternatives. For instance, upon reading *the incredibly diligent waitress*, a reader might consider that the modifier is being used to distinguish the waitress from a comparison set such as {the lazy waitress, the tired waitress, ...}. Sedivy, Tanenhaus, Chambers, and Carlson (1999) provided evidence that comprehenders often pursue this contrastive use of adjectives. In a visual world study, they had participants listen to directions like "Pick up the tall glass" while looking at a screen that had several objects, including the target (the tall glass). In the contrast condition, one of the competitor objects was another glass that was shorter, while in the no contrast condition, the competitor was an unrelated object. Among other results, they found that participants were faster to fixate the target object in the contrast condition, suggesting that they assumed the adjective was being used to draw a contrast, and that the target would therefore be the object that instantiated the contrast, e.g., the tall glass vs. the shorter one. Moreover, gradable predicates attract metrical prominence, which could also invite focus and the consideration of an alternative set.[2] Thus, if modification leads to focus in this way, then this could attract attachment of the relative clause to the modified noun.

Focus is not the only way by which modification could cause the modified noun to attract the RC. Work on retrieval interference has argued that modification can give items in memory a more elaborate representation and a higher level of activation (Arnett & Wagers, 2017; Hofmeister, 2011; Hofmeister & Vasishth, 2014). These complexity effects could also make the modified item more accessible when the parser is searching for an attachment site. Again, this account would lead to more N2 responses in the Long condition with N2 modification. Thus, I propose the Modification Attraction Hypothesis in (80) as a cover term for various theories that would predict increased attachment to a modified noun, but for different reasons.

(80) **Modification Attraction Hypothesis**: The parser is more likely to target a potential attachment site when this attachment site is modified, either because modification causes the attachment site to be interpreted as focused or because modification creates a more elaborate memory representation for the modified attachment site, making it more accessible for attachment.

[2] I thank Pranav Anand for this observation, as well as for drawing the connection to Sedivy et al.'s study.

4.4.3 The Visibility First Hypothesis

The final hypothesis acknowledges that both boundaries and focus (or modification) can influence attachment, and is an attempt to specify how the parser would weigh these two sources of evidence. According to the Visibility First Hypothesis, the parser will preferentially attach material on the basis of prosodic visibility, such that the parser prefers to target attachment sites that are more prosodically visible, where visibility is determined by the number of prosodic boundaries that intervene between the currently processed ϕ and the attachment site (Schafer, 1997). However, when two attachment sites are equally visible, then other factors like focus can tip the scales in favor of one attachment site over another.

(81) **Visibility First Hypothesis**: The parser preferentially attaches incoming material to the most visible potential attachment site, because prosodic visibility serves as a proxy for syntactic grouping. The parser weighs additional evidence, such as focus, only when two attachment sites are equally visible, because in this case visibility is insufficient for determining the most likely syntactic grouping.

4.4.4 Predictions of each account for N2 modification

The possible parses in the Short and Long condition are repeated in (82) and (83), respectively. For each condition, I have repeated the first pass prosody at the moment where the parser would encounter the relative pronoun as well as the final prosody once the full relative clause has been processed, in order to identify predictions about how prosody could influence attachment at each stage. Although visibility is important for the VFH, both potential attachment sites, *niece* and *waitress*, are in the same ϕ in the parses considered here, and so will be equally visible; I therefore do not mark visibility in these particular structures.

(82) Short Condition, First Pass

First pass: $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ ...)

Final, Option 1: $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

Final, Option 2: $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

(83) Long Condition, First Pass

First pass: $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$)

Final: $_{\phi}$(The niece$_{\omega}$ of the incredibly$_{\omega}$ diligent$_{\omega}$ waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

For the first pass prosody at the point where the parser enters the relative clause, I predict that there will not be a pre-RC boundary in the Short condition, (82), but there will be one in the Long condition, (83). For the final prosody, I predict that the presence of a pre-RC boundary will be variable in the Short condition, (82), but that there will always be one in the Long condition (83).

The RBH predicts that there will be more high attachment decisions in the Long condition relative to the Short condition, because the boundary after *waitress* will be interpreted as evidence that the following RC is not syntactically grouped with this noun. This shift to a high attachment preference should be reflected in more high attachment interpretation responses to the end-of-sentence questions. Because the pre-RC boundary already exists in the first pass prosody at the point where the RC is first encountered, it should also immediately affect attachment decisions, such that the shift to a high attachment preference in the Long condition could also be reflected in a modulation of the Ambiguity Advantage. As a reminder, the Ambiguity Advantage is the finding that there is a reading slowdown when an RC is forced to attach high compared to when it is forced to attach low or when the attachment is ambiguous. Specifically, the RBH predicts a potential change in the Ambiguity Advantage in the Long condition, such that the High Attachment penalty would be smaller and/or a Low Attachment penalty would emerge, because the parser would be more likely to pursue high attachment in this condition due to the boundary after *waitress*.

In contrast, the MAH predicts more *low* attachment in the Long condition, because modification of *waitress* will attract attachment, either by making the reader interpret this noun as focused or by making the noun more accessible. This would result in more low attachment responses to the end-of-sentence questions in the

Long condition relative to the Short condition. It may also result in a modulation of the Ambiguity Advantage, such that the Low Attachment penalty relative to the Ambiguous condition would go away in the Long condition. However, since English has a low attachment bias and this penalty already tends to be fairly small, it would be unsurprising if there is no such interaction. Note that because the MAH does not make an explicit claim about prosodic boundaries, this prediction is based on an unaugmented version of the MAH that only considers the role of modification, ignoring any role of boundaries. An augmented version of the MAH that also incorporates boundaries is also possible, and in fact, the VFH is one such way to acknowledge the influence that both sources of evidence could have on attachment.

As stated earlier, in each of the phrasings, both first pass and final, the two nouns are equally visible because they are always in the same prosodic phrase. Since neither noun is preferred based on visibility, the VFH would then predict that the baseline attachment preference in the Short condition would be determined by other factors; for instance, there may be a low attachment bias due to Late Closure or recency. In the Long condition, the VFH predicts that modification of *waitress* will attract attachment to this noun, for the same reasons as the MAH. Again, this would result in more low attachment interpretation responses and a potential modulation of the Ambiguity Advantage, such that the Low Attachment penalty could go away. In other words, for this experiment, the MAH and the VFH make the same predictions and cannot be teased apart, but I return to this question in Experiment 8, ultimately arguing that the VFH provides a better account of the data.

4.5 Experiment 7: Effects of N2 Modification

As described above, Experiment 7 tested how increasing the length of the second NP in a complex NP + RC sequence would impact the implicit prosody assigned by readers and their attachment decisions, as reflected in response times and end-of-sentence comprehension questions.[3]

4.5.1 Methods

4.5.1.1 Materials

The experiment consisted of 60 items like those in (84), manipulating the length of NP2 (Short vs. Long) and the required RC attachment (Ambiguous, Low, High). The length manipulation was achieved by adding an adverb and an adjective before N2, and RC attachment height was always disambiguated by the reflexive: half of the items were disambiguated based on gender and the other half were disambiguated based on number. The reflexive pronoun was always followed by a three-word PP, which served as a spillover region.

[3]See Appendix A for a replication of this experiment in SPR.

(84) a. The **niece** of the {incredibly diligent} **waitress** [who chided **herself** over the blunder] had just turned seventeen.

b. The son of the {incredibly diligent} **waitress** [who chided **herself** over the blunder] had just turned seventeen.

c. The **son** of the {incredibly diligent} waitress [who chided **himself** over the blunder] had just turned seventeen.

The nouns and adjectives were selected to ensure that the number of weak syllables between N1 and N2 (in the Short condition) and between Adj and N2 (in the Long condition) was always two. N1 was always monosyllabic, the adjective always ended in a SWW pattern, and N2 always started with a stressed syllable. I took this precaution because I reasoned that changes in the number of weak syllables preceding N2 could potentially affect the relative prominence of N2 between conditions for reasons other than the experimental manipulation. It is known that languages prefer to alternate strong and weak syllables, avoiding clashes and lapses, and it is plausible that after a long lapse readers could assign a particularly strong stress to N2 that could be interpreted as an accent or as greater prominence, thereby attracting attachment. Indeed, Bader (1998) showed evidence that German readers were more likely to stress a pronoun when it was preceded by a long lapse, providing support for this concern. More work on the role of lapses in silent reading is needed to determine whether a difference in the number of weak syllables preceding a noun would affect attachment preferences in the complex NP + RC structure; however, out of an abundance of caution, I controlled for this in the stimuli.

All foils were generated using Boyce et al.'s (2020) A-Maze. They were then checked manually, and foils that were deemed possible continuations were changed to be more implausible. Five items were excluded because a sentence in at least one of the conditions was ungrammatical or had an incorrect attachment height disambiguation due to experimenter error. This left 55 items in the analysis. In addition, there were 96 filler items. The complete list of experimental items is provided in Appendix E.

4.5.1.2 Participants

76 native English speakers participated in the experiment. All subjects were recruited from Prolific Academic and were paid $12 for their participation. 14 subjects were excluded because they failed to complete the Maze for at least 50% of the sentences. The data from the remaining 62 subjects were included in the analysis.

4.5.1.3 Procedure

On each trial, participants completed the Maze and then answered a comprehension question. For the experimental items, this question probed whether participants had attached the RC to N1 or N2.

4.5.2 Results

4.5.2.1 Comprehension question responses

Trials on which participants failed to complete the Maze were eliminated prior to the analysis of the comprehension question responses; a total of 767, or 22% of the 3410 trials, were removed for this reason.

The percentage of high attachment interpretations by condition is provided in Table 4.1. The percentage of high attachment responses was 45.0% and 34.4% in the Ambiguous, Short and Ambiguous, Long conditions, respectively, reflecting a tendency to provide a low attachment interpretation when either one was compatible with the sentence. When the overall sentence required a low attachment response, participants provided a high attachment response 24.2% of the time in the Short condition, and 15.2% of the time in the Long condition. For the high attachment conditions, participants provided a high attachment response 78.4% of the time in the Short condition, and 78.0% of the time in the Long condition.

ATTACHMENT	LENGTH	% High Attachment
Ambiguous	Short	45.0 (2.8)
	Long	34.3 (3.5)
Low (N2)	Short	24.2 (3.0)
	Long	15.2 (2.4)
High (N1)	Short	78.4 (2.4)
	Long	78.0 (2.5)

Table 4.1: % high attachment responses in Experiment 7

I fit a Bayesian logistic mixed effects model to the response data with the Attachment conditions Helmert coded into two contrasts: High Attachment, which compared the High condition to Ambiguous and Low, and Ambiguity, which compared the Low condition to the Ambiguous condition. The fixed effects were thus High Attachment (High vs. Ambiguous and Low), Ambiguity (Ambiguous vs. Low), Length (Short = Unmodified N2 vs. Long = Modified N2), and the interactions of Length with High Attachment and of Length with Ambiguity. The model included the maximal random effects structure.

There was a main effect of High Attachment (2.04, [1.71, 2.38]), such that participants provided more high attachment responses when the sentence required a high attachment parse in order to be grammatical, a main effect of Ambiguity (-.69,

[-.87, -.52]), such that participants provided fewer high attachment interpretations of sentences that required a low attachment parse than they did for globally ambiguous sentences, and a main effect of Length (-.25, [-.41, -.09]), such that participants provided fewer high attachment parses when N2 was lengthened via modification. There was also an interaction of High Attachment and Length (.3, [.1, .51]) reflecting the fact that the Length manipulation did not affect responses when the sentence required a high attachment parse. There was no interaction of Ambiguity and Length, indicating that modifying N2 led to a similar reduction in high attachment responses in both the Ambiguous and Low conditions. Thus, modification only influenced responses when the low attachment parse would have resulted in a grammatical sentence, i.e., in the Ambiguous and Low, but not High, conditions.

4.5.2.2 Response times

Only trials for which participants successfully completed the Maze were included in the analysis of the response times. I then excluded observations shorter than 100ms or greater than 5000ms, which was .3% of the data.

For the analysis of the complex NP region, I collapsed data across all Attachment conditions, because the primary difference between conditions in this region was the presence or absence of the adverb and adjective before N2. However, it should be noted that there were occasional differences between attachment conditions in the content of N1 and N2 (e.g., *aunt* vs. *aunts*, *king* vs. *queen*) that were collapsed over in this analysis. The mean RTs at each word in the Complex NP are shown in Figure 4.1.

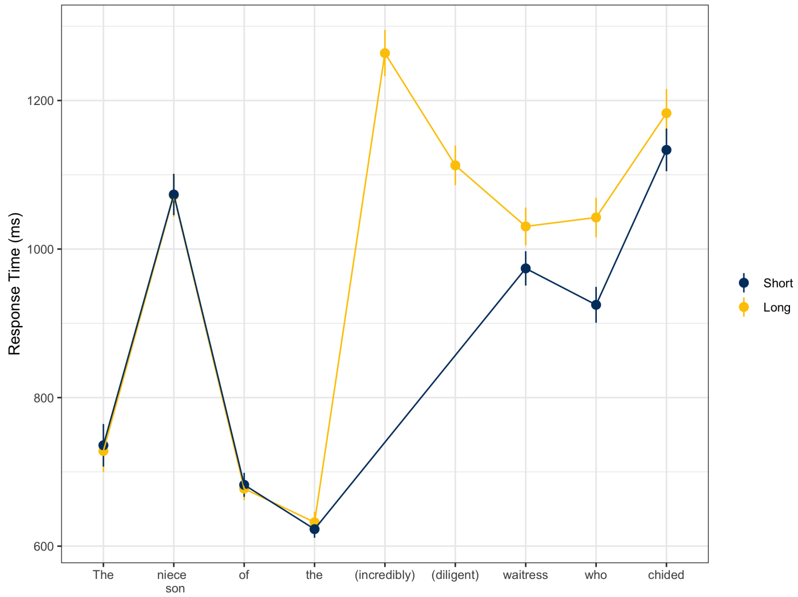


Figure 4.1: Mean response times at each word in the complex NP in Experiment 7. Error bars show a 95% confidence interval.

I fit Bayesian mixed effects models to the response times at N2, the relative pronoun *who*, and the relative clause verb *chided*, with Length as the fixed effect and the maximal random effects structure. There was a main effect of Length at N2 (25.7, [2.44, 48.53]) and at the relative pronoun (54.44, [33.63, 75.44]), such that the response times at these words were slower in the Long condition, but there was not an effect of Length at the relative clause verb (19.75, [-0.08, 39.09]). To the extent that response times reflect implicit prosodic phrasing, and under the assumption that an implicit prosodic boundary would result in longer response times, the Length effect at N2 and the relative pronoun is suggestive evidence that the modification manipulation successfully induced more and/or stronger prosodic boundaries at the end of the complex NP. I return to this point in the discussion.

Mean response times from the beginning of the sentence through the end of the relative clause are plotted in Figure 4.2. I fit Bayesian mixed effects models to the response times at the reflexive (*themselves*) and each of the two spillover words (*over* and *the*) with High Attachment (High vs. Ambiguous and Low), Ambiguity (Ambiguous vs. Low), Length (Short vs. Long), and the interactions of Length and High Attachment and of Length and Ambiguity as fixed effects and the maximal random effects structure. At the reflexive, I found a main effect of High Attachment (152.82, [107.05, 198.53]), such that there was a slowdown in the High condition, and a main effect of Ambiguity (25.28, [0.49, 50.17]), such that response times were slower in the Low condition compared to the Ambiguous condition. A main effect of High Attachment was also found at the first spillover (76.54, [43.99, 108.54]) and the second spillover (32.16, [10.38, 53.97]). No other effects were significant.

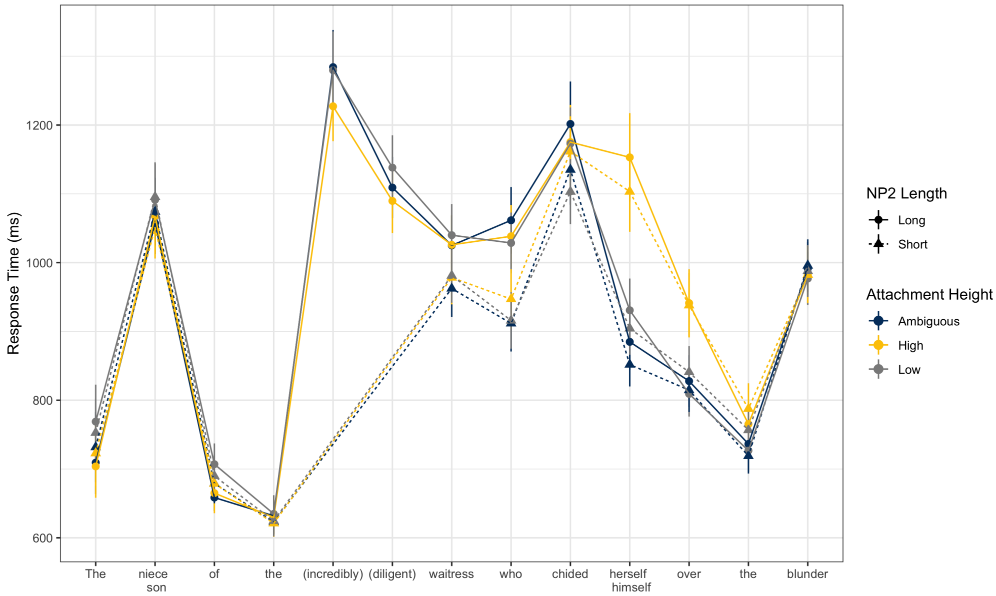


Figure 4.2: Mean response times at each word through the relative clause in Experiment 7. Error bars show a 95% confidence interval.

4.5.3 Discussion

The findings of this experiment were compatible with the MAH and VFH but not with the RBH, because lengthening NP2 led to more, not fewer, low attachment responses. Additionally, in the response time data, I replicated the Ambiguity Advantage because at the reflexive there was a main effect of High Attachment, showing a slowdown for the High condition compared to the Ambiguous and Low conditions, and a main effect of Ambiguity, such that there was a slowdown in the Low condition compared to the Ambiguous condition. However, this Ambiguity Advantage was not modulated by length. Under the RBH, I would have expected a smaller High Attachment penalty in the Long condition if the implicit boundary led to a preference for attaching high; because there was no interaction, this is additional evidence against the RBH. Since the MAH and VFH predict that the modification of N2 would only reinforce English's existing low attachment bias, the fact that Length did not modulate the Ambiguity Advantage is more compatible with these two hypotheses.

As an exploratory analysis, I was interested in whether differences in the presence or absence of an implicit prosodic boundary would be reflected in reading times. By hypothesis, I expected that readers would be more likely to assign an implicit prosodic boundary between N2 and the relative clause when N2 was modified, because longer constituents are more likely to be followed by a boundary. One might have expected that an implicit prosodic boundary would be reflected in longer reading times on either side of this boundary. For instance, if implicit prosodic boundaries are realized with syllable lengthening or a pause, then there should be longer response times to words that are hypothesized to be located at the right edge of a boundary compared to response times to words that are hypothesized to be phrase-internal. Indeed, there were longer response times at N2 and the relative pronoun in the Long condition, i.e., in the condition where I hypothesized that these two words flanked a major prosodic break. While this supports the hypothesized phrasing, I also urge caution in interpreting these results, because the field currently lacks a clear understanding of how prosodic boundaries are reflected in response times.

Previous work provides some suggestive evidence of how prosodic boundaries may manifest in the reading record. Staub (2007a) investigated whether the parser tries to attach a noun phrase as the direct object of an immediately preceding intransitive verb by comparing sentences like those in (85), manipulating both the presence of a comma between a subordinate clause and the main clause as well as whether there was intervening material (a PP or an adverb) between the verb and the main clause subject. He reasoned that there would be a reading slowdown on *the vet and his assistant* in (85a) compared to (85b) if the parser tried to attach the NP as an object of the verb, because the comma would rule out that garden path in (85b); however, under this account there would be no effect of a comma for (85c) and (85d), because the presence of a PP would rule out the direct object parse of the NP. Instead, Staub found that there was a slowdown in both no comma conditions. He argued that this slowdown was not due to the NP being incorrectly interpreted as a direct object, because the slowdown also appeared in (85c); rather, he suggested that this slowdown could be related to implicit prosody. According to this account, the absence of a comma where a prosodic boundary would be expected was disruptive, and readers responded as they would to a prosodic violation by slowing down. Thus, this experiment suggests that reading times are increased on the word *following* the location where an implicit boundary is assigned.

(85) a. When the dog arrived the vet and his assistant went home.

b. When the dog arrived, the vet and his assistant went home.

c. When the dog arrived at the clinic the vet and his assistant went home.

d. When the dog arrived at the clinic, the vet and his assistant went home.

Staub (2007a) cited several other experiments with similar findings. Adams, Clifton, and Mitchell (1998) found longer reading times at *the veterinarian* in (86a), when readers know it must be the beginning of a new clause, compared to (86b), where they think it is a direct object of the subordinate clause. Staub also cited an unpublished study, Adams, Clifton, and Mitchell (1991), which found that *her assistant* was read more slowly in the version of (87) without a comma. Both of these findings support the idea that the absence of a comma where a prosodic boundary is expected leads to a slowdown.

(86) a. After the dog scratched pathetically the veterinarian took off the muzzle.

b. After the dog scratched the veterinarian took off the muzzle.

(87) While the announcer read the names(,) her assistant checked the next item on the list.

An alternative possibility is that this slowdown is not disruptive, but reflects the assignment of the implicit boundary. In line with the principle of Prosodic Late Closure from Chapter 3, the parser might have a tendency to wait to assign a prosodic boundary until it has unambiguous evidence that it is at a phrase (or clause) edge. In the no comma version of (87), the parser may suspect that it is at a phrase edge, but would also know that this phrase could be followed by additional clause-internal phrases, such as a PP like *over the intercom*. Once it sees *her assistant*, the parser would have unambiguous evidence that the subordinate clause was complete, and would now assign the boundary between the subordinate clause and the matrix clause while fixated on the right side of the boundary, resulting in slower reading times on that region. These findings would also be consistent with a world in which prosodic parsing generally lags behind syntactic parsing, such that the assignment of implicit prosodic boundaries shows up on the region to the right of said boundary in reading time measures. Setting aside whether or not this slowdown represents a disruption, the main takeaway from Staub (2007a) is that there is potentially a reading slowdown on the word to the right of the location where an implicit boundary is hypothesized to be assigned. This might grant additional confidence in interpreting the slowdown on N2 and the relative pronoun as reflecting a boundary.

However, there are also reasons to be skeptical of correlating response times with the existence of implicit boundaries. In overt speech, boundaries can be conveyed by means other than syllable lengthening and pauses, including boundary tones and a pitch reset at the start of a new phrase (Wagner & Watson, 2010). Under the assumption that syllable lengthening and pauses would be reflected in a reading slowdown, but these other cues would not, then a reader could have represented an implicit boundary without there being a reading slowdown in that position. Note, however, that this assumption about how different implicit prosodic features may be reflected in reading times is merely an intuition in need of confirmation in future research.

Previous work also suggests that the connection between reading measures, the amount of prosodic structure, and the duration of the same material in spoken language is not always straightforward. Ashby and Clifton Jr (2005) found that words with two stressed syllables took longer to read and were more likely to be refixated than words with one stressed syllable. However, within single-stressed one and two-syllable words, they did not find any differences in mean first fixation durations or mean gaze durations, even though the two-syllable words had longer naming durations in a production experiment. Thus, while the longer response times on N2 and the relative pronoun in the Long condition could plausibly reflect implicit prosodic phrasing, and although this interpretation is compatible with my claim that participants were more likely to assign a boundary in this position, I contend that further work is necessary in order to confidently draw conclusions about implicit prosody from response times in this task.

Related to this issue, there is an alternative interpretation of the findings, according to which the Repellent Boundaries Hypothesis is right but my assumptions about the implicit prosody are wrong: perhaps increasing the length of N2 did not actually increase the likelihood of placing a boundary in this position. However, this account would go against much previous work showing that boundaries are more frequent after long constituents, and it would also still need to explain why there was a higher rate of low attachment in the Long condition. That being said, future work should probe the prosodic structures that participants assign. For instance, a rapid prosodic transcription task in which participants are trained to mark where they place boundaries in their inner speech while reading could shed light on how participants parsed the stimuli; see also Bishop (2020). Such a task could also try to correlate the reported phrasings with the interpretations that participants provide, which would help strengthen the argument in favor of the MAH and VFH by establishing that participants indeed provide more low attachment responses when N2 is modified, even when they place a boundary between N2 and the RC. A production study is another option to see which phrasings people most frequently provide. There are some limitations to these approaches, however. Overt productions might not always reflect implicit prosody (Jun, 2010), and this is particularly true when one considers that the assignment of implicit prosody involves both a first pass prosody and potential subsequent revisions to this prosody, whereas the prosody in production experiments reflects only the final prosody that the speaker settled on, possibly after rehearsal of the target sentence.

To summarize, the main takeaway from Experiment 7 is that modification of the second noun led to more low attachment responses, and there was suggestive evidence from response times for a prosodic boundary in the hypothesized position between the complex NP and the RC, though this requires further research. The effects on attachment are consistent with the predictions of the MAH and the VFH, which state that modification should attract attachment, but not the RBH, which states that the implicit prosodic boundary preceding the RC should encourage non-local attachment. In order to tease apart the MAH and the VFH, I test the effect of modifying N1 instead of N2 in Experiment 8.

4.6 Experiment 8: Effects of N1 Modification

The findings from Experiment 7 are consistent with the MAH and the VFH. However, the two accounts diverge in their predictions for what happens when N1 is modified. First, consider the possible prosodic phrasings when the first noun *niece* is modified, as in (88). As before, for the Short condition, I predict that there will variably be a prosodic break between the second noun *waitress* and the RC, as in (89a,b). Predictions for the Long condition are less straightforward. While I expect that modifying the first noun will increase the likelihood of a prosodic boundary in general, there are two plausible places where this boundary could go: after the first noun, *niece*, as in (90a), or after the entire complex NP, as in (90b). I expect that both parses in (90) are likely, and so there may be variable phrasing in the Long condition.

(88) a. The niece$_{\text{N1}}$ of the waitress$_{\text{N2}}$ who chided$_{\text{V}}$ herself$_{\text{Refl}}$ over the blunder$_{\text{PP}}$...

b. The incredibly$_{\text{Adv}}$ diligent$_{\text{Adj}}$ niece$_{\text{N1}}$ of the waitress$_{\text{N2}}$ who chided$_{\text{V}}$ herself$_{\text{Refl}}$ over the blunder$_{\text{PP}}$...

(89) Short condition

a. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

b. $_{\phi}$(The niece$_{\omega}$ of the waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

(90) Long condition

a. $_{\phi}$(The incredibly$_{\omega}$ diligent$_{\omega}$ niece$_{\omega}$) $_{\phi}$(of the **waitress**$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

b. $_{\phi}$(The incredibly$_{\omega}$ diligent$_{\omega}$ niece$_{\omega}$ of the waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

Crucially, however, the MAH and the VFH make different predictions about attachment for the phrasings in the Long condition (90). Under the MAH, the phrasing will not matter: regardless of where the prosodic boundary is placed, modification of the first noun, *niece*, will lead to more high attachment in the Long condition because modification makes this noun focused and/or more accessible. This change in attachment preferences will be reflected in the question responses and potentially in an attenuated or reversed High Attachment penalty in the response times at the reflexive. Under the VFH, there will be more low attachment interpretation responses in (90a) even though the first noun is modified. In (90a), the second noun, *waitress* is more visible than the first noun, *niece*, because there is an additional prosodic boundary separating the first noun from the RC; this difference in visibility is represented by putting the more visible noun, *waitress*, in boldface. Visibility takes precedence over modification in attachment under the VFH, leading to more low attachment responses. But in (90b), both nouns are equally visible because no prosodic boundary intervenes between them. However, only the first noun, *niece*, is modified. Under the VFH, modification affects attachment preferences when two attachment sites are equally visible, meaning that there will be more high attachment responses in (90b). This means that the effect on attachment in the experiment will largely depend on the relative probability of these two parses, such that the VFH could accommodate a shift to more N1 responses or a shift to more N2 responses. I will show that there is a slight uptick in N2 responses when N1 is modified, which is only consistent with the VFH but not with the MAH.

4.6.1 Methods

4.6.1.1 Materials

The same materials from Experiment 7 were used, with the exception that N1 was now modified instead of N2, as in (91). The errors in the five items that had to be excluded in Experiment 7 were corrected, such that all 60 items were included in the analysis of Experiment 8.

(91) a. The {incredibly diligent} **niece** of the **waitress** [who chided **herself** over the blunder] had just turned seventeen.

b. The {incredibly diligent} son of the **waitress** [who chided **herself** over the blunder] had just turned seventeen.

c. The {incredibly diligent} **son** of the waitress [who chided **himself** over the blunder] had just turned seventeen.

4.6.1.2 Participants

71 native English speakers participated in the experiment. 30 subjects were undergraduate students at the University of California, Santa Cruz who completed the study for course credit, and 42 were recruited from Prolific Academic and were paid $12 for participating. 11 subjects were excluded due to poor performance on the Maze, again defined as successfully completing fewer than 50% of trials, leaving data from 60 subjects for analysis.

4.6.1.3 Procedure

The same procedure from Experiment 7 was used.

4.6.2 Results

4.6.2.1 Comprehension question responses

I eliminated responses from the 797 trials (22%) on which participants failed to complete the Maze prior to analyzing the comprehension question responses. The percentage of high attachment interpretations by condition is provided in Table 4.2.

ATTACHMENT	LENGTH	% High Attachment
Ambiguous	Short	53.3 (3.4)
	Long	45.7 (3.0)
Low (N2)	Short	31.3 (2.6)
	Long	23.7 (2.8)
High (N1)	Short	73.9 (3.2)
	Long	76.5 (2.9)

Table 4.2: % high attachment responses in Experiment 8

As in the previous experiment, I fit a Bayesian logistic mixed effects model to the response data with High Attachment (High vs. Ambiguous and Low), Ambiguity (Ambiguous vs. Low), Length (Short = Unmodified N1 vs. Long = Modified N1), and the interactions of High Attachment and Length and of Ambiguity and Length as fixed effects, and the maximal random effects structure. There was a main effect of High Attachment (1.49, [1.18, 1.80]), such that more high attachment interpretations were provided in the High condition compared to Ambiguous and Low, and a main effect of Ambiguity (-.64, [-.81, -.47]), such that fewer high attachment interpretations were provided in the Low condition compared to the Ambiguous condition. Although there was no main effect of Length (-.09, [-.21, .03]), there was a significant interaction of High Attachment and Length (.25, [.08, .42]), reflecting that modification of N1 led to more low attachment responses in the Ambiguous and Low conditions but not in the High condition. To confirm that modification of N1 led to more N2 interpretation responses in the globally ambiguous sentences, a separate model was fit to only the responses to the Ambiguous conditions, with Length as the fixed effect and the maximal random effects structure. There was a significant effect of Length (-.18, [-.35, -.02]), such that participants were more likely to provide low attachment responses when the first noun was modified. This finding goes against the Modification Attraction Hypothesis, and I return to this point in the discussion.

4.6.2.2 Response times

Observations shorter than 100ms or greater than 5000ms were excluded from the analysis, which was .2% of the data. As in Experiment 7, I collapsed across all Attachment conditions for analysis of the complex NP region, and the mean RTs for each word in the complex NP are plotted in Figure 4.3. Bayesian mixed effects models were fit to the response times at N1, the preposition *of*, Det2, N2, the relative pronoun *who*, and the relative clause verb *chided*, with Length as the fixed effect and the maximal random effects structure.

There was a main effect of Length at N1 (-84.92, [-106.35, -63.32]), such that response times were faster when N1 was modified. An effect of length in the opposite direction was found at the following preposition (12.98, [.93, 25.07]), with slower responses when N1 was modified. At Det2, response times were once again faster in the Long condition (-9.53, [-17.7, -1.38]). There was no main effect of Length at N2 (-2.98, [-18.44, 12.58]), but there was a main effect of Length at the relative pronoun (28.85, [11.44, 46.15]) and at the relative clause verb (21.35, [3.88, 38.97]), such that response times at these words were longer when N1 was modified. I consider possible reasons for these effects in the discussion.

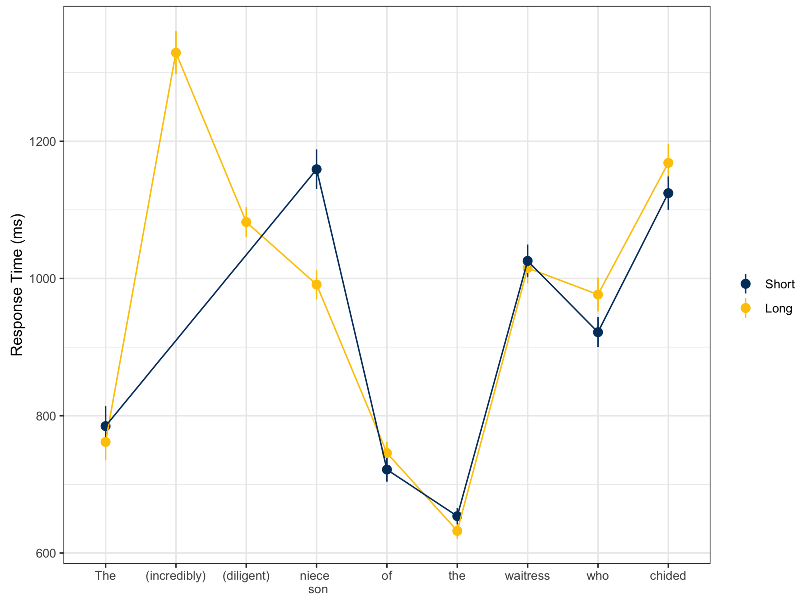


Figure 4.3: Mean response times at each word in the complex NP in Experiment 8. Error bars show a 95% confidence interval.

Mean response times through the end of the relative clause are shown in Figure 4.4. As before, I fit Bayesian mixed effects models to the response times at the reflexive (*themselves*) and each of the two spillover words (*over* and *the*) with High Attachment (High vs. Ambiguous and Low), Ambiguity (Ambiguous vs. Low), Length (Short = Unmodified N1 vs. Long = Modified N1), and the interactions of Length and High Attachment and of Length and Ambiguity as fixed effects and the maximal random effects structure. At the reflexive, there was a main effect of High Attachment (130.23, [85.42, 174.31]), such that response times were slower when the reflexive disambiguated to a high attachment parse, and a main effect of Ambiguity (24.33, [2.42, 46.13]), such that response times were slower when the reflexive disambiguated to a low attachment parse compared to the Ambiguous conditions. There were no other significant effects at the reflexive. The significant High Attachment effect was also found at the first spillover word (42.21, [13.18, 71.53]) but not at the second spillover word (14.31, [-5.50, 34.30]). No other effects were significant.

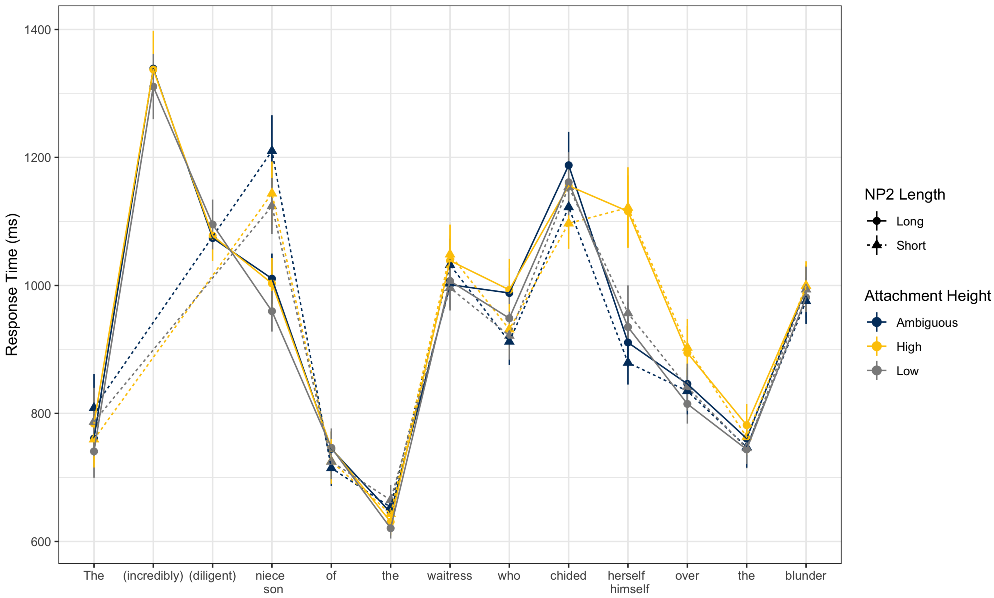


Figure 4.4: Mean response times at each word through the relative clause in Experiment 8. Error bars show a 95% confidence interval.

4.6.3 Discussion

I found that modifying N1 led to more N2 interpretation responses. I also replicated the Ambiguity Advantage again: for the response times at the reflexive, there was a main effect of High Attachment, such that there was a reading slowdown in the High condition compared to the Ambiguous and Low conditions, and a main effect of Ambiguity, such that there was a reading slowdown in the Low condition compared to the Ambiguous condition. However, as in Experiment 7, the Ambiguity Advantage was not modulated by Length. Both the interpretation responses and the response times at the reflexive are inconsistent with the Modification Attraction Hypothesis, which predicted more high attachment responses and a smaller High Attachment penalty in the Long condition due to the fact that the length was achieved by modifying N1.

However, these findings are consistent with the Visibility First Hypothesis, which states that boundaries gate modification, such that modification only influences attachment preferences when two attachment sites are equally visible, i.e., separated from the currently processed phrase by the same amount of boundaries. Under this view, modification of N1 led to two possible prosodic parses, repeated in (92). In (92a), the second noun, *waitress*, is more visible than the first noun, *niece*, so there is an overall preference for low attachment. In (92b), the two nouns are equally visible; with this parse, modification of *niece* would favor high attachment.

(92)

a. $_{\phi}$(The incredibly$_{\omega}$ diligent$_{\omega}$ niece$_{\omega}$) $_{\phi}$(of the **waitress**$_{\omega}$ who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

b. $_{\phi}$(The incredibly$_{\omega}$ diligent$_{\omega}$ niece$_{\omega}$ of the waitress$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$

over the blunder$_\omega$)

I have suggested both of the phrasings in (92) were the most likely parses, and that the parser may have variably chosen between these two phrasings. Based on this study, I cannot determine the relative frequency of these two phrasings within this experiment, i.e., the rate at which the parser assigned the structure in (92a) vs. (92b). At first glance, the shift toward more low attachment interpretation responses in the Long condition might suggest that the parse in (92a), in which the second noun is more visible, was more common within the experiment than the parse in (92b), because only the former would favor low attachment under the VFH. However, this need not be the case: in principle, it would be possible to obtain the same percentage of high attachment responses even if the parser pursued the phrasing in (92b) on a greater number of trials, provided that modification of the first noun only introduced a weak preference for high attachment. For the sake of argument, and without external motivation, assume that the parser pursued (92a) 30% of the time and (92b) 70% of the time. Next, assume that (92a), with greater visibility of the second noun, *waitress*, results in high attachment only 40% of the time, but that the modification of the first noun, *niece*, in (92b) results in a modest increase to 50% high attachment. This would result in an overall high attachment response rate of .7(40) + .3(50) = 47% high attachment responses, similar to the 45.7% high attachment interpretation responses. I do not know the actual rate at which the parser pursued one phrasing or the other in this experiment, nor do I know the magnitude of the influence of visibility or modification on the rate at which the parser pursues low versus high attachment. Since it is in principle possible that the observed shift to low attachment responses in the Long condition could have been achieved even if the parser pursued the phrasing in (92b) on a greater number of trials, I cannot draw any conclusions about the relative frequency within the experiment of the two phrasings based on the interpretation data.

It is also unclear what the reading data can tell us about the preferred phrasing in the Long condition. On the one hand, the response time slowdown at the first noun might suggest the parse in (92). However, this noun was the second word in the Short condition and the fourth word in the Long condition. I have noticed that response times at the second word in the Maze are often inflated; a likely contributor is that the preceding context is not yet very constraining after just one word. Thus, the longer RTs in the Short condition, which suggested a boundary at N1 in the Short condition but not in the Long condition, could be artifactual. This is potentially supported by the fact that there was a slowdown in the Long condition at the preposition immediately following N1. If, as discussed in Experiment 7, response times tend to be slower on the word to the right of an implicit boundary, then this would actually suggest that a boundary was placed after N1 more frequently in the Long condition compared to the Short condition, and that (92a) was more common. However, there was also a slowdown at the relative pronoun and the relative clause verb, which might suggest a prosodic boundary after N2, in support of the alternative parse, (92b).

One possibility is that some of these effects are spurious. It could also be the case that the evidence for boundaries in different positions reflects the fact that the parser variably pursued each parse. Yet another possibility is that participants actually pursued a parse in which there was a boundary after the first noun, *niece*, *and* after the second noun, *waitress*, as in (93).

(93) $_{\phi}$(The incredibly$_{\omega}$ diligent$_{\omega}$ niece$_{\omega}$) $_{\phi}$(of the **waitress**$_{\omega}$) $_{\phi}$(who chided$_{\omega}$ herself$_{\omega}$ over the blunder$_{\omega}$)

The phrasing in (93) would violate BINMIN, the constraint which requires each phrase to contain at least two words, because the phrase containing *waitress* only contains one prosodic word, but it is arguably more balanced because the sequence of 3 words + 1 word + 3 words, *(the incredibly diligent niece) (of the waitress) (who chided herself over the blunder)*, is more symmetrical than the parses in (92a), which has a sequence of 3 words + 4 words, *(the incredibly diligent niece) (of the waitress who chided herself over the blunder)*, and (92b), with 4 words + 3 words, *(the incredibly diligent niece of the waitress) (who chided herself over the blunder)*. This phrasing in (93) would also favor low attachment responses, because the second noun, *waitress*, which is only separated from the RC by one prosodic boundary, would be more visible than the first noun, *niece*, which is separated from the RC by two boundaries. Overall, it is unclear what to make of these response time data, and as discussed in Experiment 7, there is already reason to be skeptical about drawing inferences about implicit prosody from response times without further research. Again, a rapid prosodic transcription task or a production study would be useful to see which phrasings tend to be most frequent, especially if these phrasings can then be correlated with interpretation data. However, as I have frequently noted, there are potential differences between first pass and final parses, such that results from these studies may be more informative about the final prosody than about the incremental assignment of prosody.

The important conclusion from this experiment is that the responses cannot be driven *solely* by modification, without any reference to phrasing: otherwise, regardless of where boundaries were or were not placed in each condition, the presence of modification in the Long condition but not in the Short condition would favor more high attachment responses in the former. In contrast, the Visibility First Hypothesis is an explicit proposal of how prosodic visibility would gate the use of modification, allowing both sources of evidence to contribute to attachment preferences while predicting that prosodic visibility should be the primary determinant. I return to ways the Visibility First Hypothesis could be tested in future work in the General Discussion. Next, I report a follow-up experiment that replicates the main findings of Experiments 7 and 8 but with the complex NP in object position, providing convergent evidence for the VFH in a different configuration.

4.7 Experiment 9: Effects of NP Modification and RC length

The aim of Experiment 9 was to test the generality of the findings of Experiments 7 and 8 by replicating the effects of N1 and N2 modification on attachment preferences when the complex NP was in object position, as in (94). The RC length was also manipulated, in order to provide a test of the effects of modification in a broader range of prosodic structures, as I will lay out in the predictions below.

(94) a. The doctor met the son of the colonel who died.

b. The doctor met the son of the colonel who tragically died of a stroke.

c. The doctor met the relatively famous son of the colonel who died.

d. The doctor met the relatively famous son of the colonel who tragically died of a stroke.

e. The doctor met the son of the relatively famous colonel who died.

f. The doctor met the son of the relatively famous colonel who tragically died of a stroke.

The other reasons to run this experiment were primarily methodological. In the replication of Hemforth et al. (2015) in Experiment 5, I found that the RC length effect was apparently neutralized when the complex NP was in subject position in the Maze, such that RC length did not have an effect on the rate of high attachment responses. This raised concerns that length effects may not be as strong in subject position (at least in the Maze), which could have driven the greater rate of high attachment interpretation rates in Experiment 7, if participants discounted implicit boundaries and relied more on modification and/or focus because of some property specific to subject position. This concern was partially allayed by the results of Experiment 8, where modifying the first noun did *not* lead to a higher rate of high attachment responses; the increase in low attachment responses would have been unexpected if participants had based their responses solely on which noun was modified. However, to grant additional confidence in the VFH, I decided to replicate the findings in a novel configuration. Finally, in order to accommodate the additional conditions of the RC length manipulation, I removed the Attachment Height factor from this experiment; all experimental items were globally ambiguous.

4.7.1 Predicted prosodies

Although I do not walk through the step-by-step assignment of the prosodic structures for this experiment, I briefly describe the predicted preferred final prosodic structures based on the pressures of syntax-prosody correspondence, size, and balance. After going through all of the predicted structures, I discuss the competing predictions of the VFH and the MAH.

For the No Modification, Short RC condition in (95), I predict that two parses are likely: in (95a), there is no prosodic boundary, while in (95b), a boundary has been placed after the first noun *son* to achieve a balanced parse. Crucially, when the RC only contains one word, I assume that it cannot stand in a ϕ on its own and will always phrase with the preceding noun, *colonel*, following Fodor (2002a). For the No Modification, Long RC condition in (96), I predict that the long RC will now stand alone in its own ϕ, and a boundary will appear after the second noun (*colonel*), again following Fodor (2002a).

(95) No Modification, Short RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$ who died$_\omega$)

b. (The doctor$_\omega$ met$_\omega$ the son$_\omega$) (of the **colonel**$_\omega$ who died$_\omega$)

(96) No Modification, Long RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$) (who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

For both Noun 1 Modification conditions in (97) and (98), I predict that the break will come after the modified noun, *son*, following the same logic from Experiments 7 and 8, which stated that long constituents are more likely to be followed by a prosodic break (Breen et al., 2011; Hwang & Schafer, 2009; Watson & Gibson, 2004b). In the Short condition, a break in only this position could result in an imbalanced parse, as in (97a), which has a ϕ containing five words followed by a ϕ containing two words. Thus, the parse in (97b), which creates a more balanced parse by adding another boundary between the verb and the complex NP, is also likely and possibly preferred. The important point for deriving predictions is that in either case I expect a prosodic boundary after the noun *son* in this condition. In the Noun 1 Modification, Long RC condition in (98), there may also be a parse in which the boundary comes after the second noun, *colonel*, instead of the first noun, *son*; however, this parse is rather imbalanced, because the first ϕ contains six words, while the second ϕ contains three. As such, I assume that this parse is dispreferred and that the phrasing in (98a) is the frontrunner.

(97) Noun 1 Modification, Short RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$ who died$_\omega$)

b. $_\phi$(The doctor$_\omega$ met$_\omega$) $_\phi$(the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$

who died$_\omega$)

(98) Noun 1 Modification, Long RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$ who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

b. ? $_\phi$(The doctor$_\omega$ met$_\omega$ the relatively$_\omega$ famous$_\omega$ son$_\omega$ of the colonel$_\omega$) $_\phi$(who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

For the Noun 2 Modification, Short RC condition, I predict that the first noun *son* will phrase with the preceding subject and verb, while the modified second noun *colonel* will phrase with the RC; this results in a relatively balanced parse, because the first ϕ in (99a) contains three prosodic words while the second ϕ contains four. For the Noun 2 Modification, Long RC condition, I predict that the first noun *son* will phrase with the preceding subject and verb; the PP containing the modified second noun will constitute its own ϕ, and the RC will also be in its own ϕ. This results in the balanced parse in (100a) in which each ϕ contains three prosodic words.

(99) Noun 2 Modification, Short RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the son$_\omega$) $_\phi$(of the relatively$_\omega$ famous$_\omega$ **colonel**$_\omega$ who died$_\omega$)

(100) Noun 2 Modification, Long RC

a. $_{\phi}$(The doctor$_{\omega}$ met$_{\omega}$ the son$_{\omega}$) $_{\phi}$(of the relatively$_{\omega}$ famous$_{\omega}$ **colonel**$_{\omega}$) $_{\phi}$(who tragically$_{\omega}$ died$_{\omega}$ of a stroke$_{\omega}$)

4.7.2 Predicted effects on interpretation

The first prediction involves the RC length manipulation. Within each level of Modification, the Long RC condition is always more likely to have a pre-RC boundary, which should favor more high attachment responses relative to the corresponding Short RC condition, in which the second noun, *colonel*, phrases with the RC; as before, this pre-RC boundary should favor more non-local attachments compared to the Short RC condition with the same amount of modification, replicating the classic Fodor effect.

Next, compare the No Modification, Short RC condition to the Noun 1 Modification, Short RC condition, repeated in (101) and (102). In the No Modification, Short RC condition, the two nouns are sometimes equally visible, as in (101a), but the second noun, *colonel*, is more visible in (101b). In the Noun 1 Modification, Short RC condition, the second noun, *colonel*, is always more visible, because there always an additional boundary intervening between *son* and *colonel*.

The MAH would predict that there should be more high attachment responses in the Noun 1 Modification condition, because modification attracts attachment. The VFH would predict that there should be more low attachment responses in the Noun

1 Modification condition, because the two nouns *son* and *colonel* are sometimes equally visible in the No Modification condition (101), whereas the second noun *colonel* is always more visible than the first noun *son* in the phrasings in (102), such that visibility will always encourage a stronger low attachment preference in the Noun 1 Modification condition.

(101) No Modification, Short RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$ who died$_\omega$)

b. (The doctor$_\omega$ met$_\omega$ the son$_\omega$) (of the **colonel**$_\omega$ who died$_\omega$)

(102) Noun 1 Modification, Short RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$ who died$_\omega$)

b. $_\phi$(The doctor$_\omega$ met$_\omega$) $_\phi$(the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$ who died$_\omega$)

In the comparison of the No Modification, Long RC condition and the Noun 1 Modification, Long RC condition, repeated in (103) and (104), there is once again a difference in visibility. In the No Modification, Long RC condition, (103), the two potential attachment sites, *son* and *colonel*, are equally visible because they are in the same ϕ. In the Noun 1 Modification, Long RC condition, the second noun *colonel* is more visible than the first noun *son*, due to the prosodic boundary that intervenes between them. The MAH would again predict more high attachment responses in the Noun 1 Modification, Long RC condition, while the VFH would predict more low attachment responses in the Noun 1 Modification, Long RC condition because the second noun *colonel* is always more visible than the modified noun *son*, such that modification should not shift attachment preferences.

(103) No Modification, Long RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$) (who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

(104) Noun 1 Modification, Long RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the relatively$_\omega$ famous$_\omega$ son$_\omega$) $_\phi$(of the **colonel**$_\omega$ who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

For the comparison of the No Modification, Short RC condition in (105) and the Noun 2 Modification, Short RC condition in (106), there is a difference in visibility: in the No Modification condition, (105), the second noun *colonel* is only sometimes more visible than the first noun *son*. However, in the Noun 2 Modification condition, *colonel* is always more visible than the first noun *son*, because they never appear in the same prosodic phrase. Thus, the MAH predicts more low attachment responses in the Noun 2 Modification condition, because modification should attract attachment, and the VFH predicts more low attachment responses in the Noun 2 Modification condition, because the second noun is always more visible than the first.

(105) No Modification, Short RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$ who died$_\omega$)

b. (The doctor$_\omega$ met$_\omega$ the son$_\omega$) (of the **colonel**$_\omega$ who died$_\omega$)

(106) Noun 2 Modification, Short RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the son$_\omega$) $_\phi$(of the relatively$_\omega$ famous$_\omega$ **colonel**$_\omega$ who died$_\omega$)

For the final comparison, consider the No Modification, Long RC condition in (107) and the Noun 2 Modification, Long RC condition in (108). In the No Modification, Long RC condition, the first noun *son* and the second noun *colonel* are equally visible, because they are in the same prosodic phrase. In the Noun 2 Modification, Long RC condition, the second noun, *colonel*, is more visible than the first noun, *son*, because a boundary intervenes between them. The MAH predicts a greater rate of low attachment in the Noun 2 Modification condition, because modification should attract attachment via focus or greater accessibility of the modified noun. The VFH also predicts a greater rate of low attachment in the Noun 2 Modification condition, because the second noun is always more visible than the first.

(107) No Modification, Long RC

a. (The doctor$_\omega$ met$_\omega$ the son$_\omega$ of the colonel$_\omega$) (who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

(108) Noun 2 Modification, Long RC

a. $_\phi$(The doctor$_\omega$ met$_\omega$ the son$_\omega$) $_\phi$(of the relatively$_\omega$ famous$_\omega$ **colonel**$_\omega$) $_\phi$(who tragically$_\omega$ died$_\omega$ of a stroke$_\omega$)

I have now walked through predictions for the essential comparisons between the No Modification baseline conditions and the conditions in which either of the nouns is modified. To reiterate, I expect that within each modification condition, a longer RC will lead to more high attachment responses, replicating the classic RC length effect (Fodor, 1998, 2002a, 2002b; Hemforth et al., 2015), because it encourages the placement of a prosodic boundary between the second noun and the relative clause. Regardless of RC length, the MAH predicts that modification of the first noun should lead to more high attachment responses, while the modification of the second noun should lead to more low attachment responses, because modification attracts attachment to the modified noun via focus or increased accessibility. The VFH predicts that modification of the first noun should lead to more low attachment responses in the Noun 1 Modification conditions compared to the No Modification conditions, because the extra length leads to a prosodic boundary after the first noun *son*, rendering the second noun more visible. For the modification of the second noun, the VFH predicts that there will be more low attachment responses in the Noun 2 Modification conditions compared to the No Modification conditions, because modification leads to a consistent boundary before the second noun that renders it more visible than the first noun in the Noun 2 Modification conditions, whereas the second noun is only sometimes more visible than the first noun in the No Modification conditions.

4.7.3 Methods

4.7.3.1 Materials

The experiment consisted of 48 items like those in (94), repeated in (109), manipulating the presence of modification (None vs. N1 vs. N2) and the length of the relative clause (Short vs. Long). The complex NP always appeared in object position, in order to test whether the findings of Experiments 7 and 8, in which the complex NP was in subject position, would generalize to this new position. As in Experiments 7 and 8, modification was achieved by introducing a two-word phrase consisting of an adverb and an adjective. 40 of the items were adapted from Experiments 5 and 6 (the replication of Hemforth et al. (2015)), while the remaining 8 were written for this experiment. The full list of experimental items is provided in Appendix G.

(109) a. The doctor met the son of the colonel who died.

b. The doctor met the son of the colonel who tragically died of a stroke.

c. The doctor met the relatively famous son of the colonel who died.

d. The doctor met the relatively famous son of the colonel who tragically died of a stroke.

e. The doctor met the son of the relatively famous colonel who died.

f. The doctor met the son of the relatively famous colonel who tragically died of a stroke.

4.7.3.2 Participants

67 native English speakers participated in the experiment. Participants were undergraduate students at the University of California, Santa Cruz who completed the study for course credit. 3 subjects were excluded due to poor performance on the Maze, again defined as successfully completing fewer than 50% of trials, leaving data from 64 subjects for analysis.

4.7.3.3 Procedure

The same procedure from Experiments 7 and 8 was used.

4.7.4 Results

4.7.4.1 Comprehension question responses

Responses from the 535 trials (17%) on which participants failed to complete the Maze were eliminated prior to analysis of the comprehension question responses. The percentage of high attachment interpretations by condition is provided in Table 4.3.

MODIFICATION	RC LENGTH	% High Attachment
None	Short	30.9 (3.6)
	Long	42.0 (3.2)
Noun 1	Short	25.6 (3.5)
	Long	37.6 (3.5)
Noun 2	Short	13.8 (2.5)
	Long	23.1 (2.7)

Table 4.3: % high attachment responses in Experiment 9

I fit a Bayesian logistic mixed effects model to the response data with treatment coding such that the Noun 1 Modification conditions and the Noun 2 Modification conditions were each compared separately to the No Modification conditions, which served as the baseline. Thus, the fixed effects were Noun 1 Modification (None vs. Noun 1), Noun 2 Modification (None vs. Noun 2), RC Length (Short RC vs. Long

RC), and the interaction of Length and Noun 1 Modification and of Length and Noun 2 Modification as fixed effects, and the maximal random effects structure. There was a main effect of Length (.34, [.13, .56]), such that participants provided more high attachment responses in the Long RC condition, a main effect of Noun 1 Modification (-.43, [-.8, -.1]), such that participants provided more low attachment interpretations when the first noun was modified than when neither noun was, and a main effect of Noun 2 Modification, (-1.41, [-1.85, -1.03]), such that participants provided more low attachment interpretations when the second noun was modified than when neither noun was. There were no significant interactions. These findings pattern with Experiments 7 and 8, favoring the VFH over the MAH; I return to this in the discussion.

4.7.4.2 Response times

Observations shorter than 100ms or greater than 5000ms were excluded from the analysis, which was .5% of the data. Because the material in the Short RC and Long RC conditions was identical through the relative pronoun, I collapsed across the RC Length conditions for analysis of the response times in the complex NP region. The mean response times are plotted in Figure 4.5. Bayesian mixed effects models were fit to the response times at the first noun *son*, the preposition *of*, the determiner before the second noun, the second noun *colonel*, and the relative pronoun *who*, with Noun 1 Modification (None vs. Noun 1) and Noun 2 Modification (None vs. Noun 2) as fixed effects, and the maximal random effects structure.

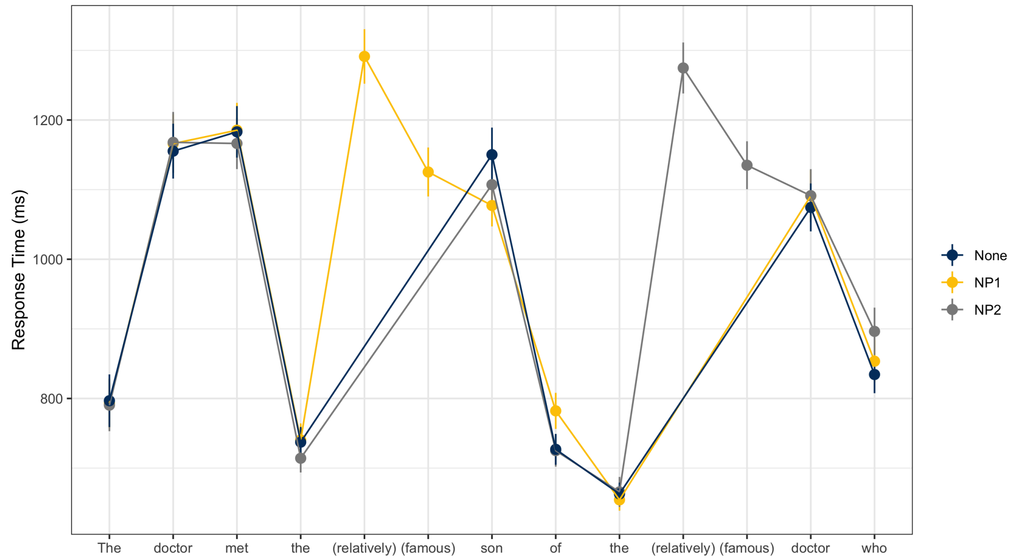


Figure 4.5: Mean response times at each word in the complex NP in Experiment 9. Error bars show a 95% confidence interval.

At the first noun, there was a main effect of Noun 1 Modification (-77, [-126, -28]) , such that response times were faster on the first noun when it was modified. This replicates the finding from Experiment 8, and suggests that if the parser placed a boundary after the modified N1 as predicted, then it did not show up on the noun itself. There was also a main effect of Noun 2 Modification (-50, [-100, -.2]), such that response times were shorter when the second noun was modified, but because the Noun 2 and No Modification conditions were identical at this position I conclude that this effect is spurious.

At the preposition, there was a main effect of Noun 1 Modification (53, [18, 87]), such that response times were longer when the first noun was modified compared to when neither noun was. This potentially reflects the implicit prosodic boundary that I predicted participants would place after the modified first noun, provided that the assignment of an implicit prosodic boundary is reflected in longer response times on the word that follows the boundary, as discussed in Experiment 7.

There were no significant effects at the determiner or at the second noun. At the relative pronoun, there was a main effect of Noun 2 Modification (58, [9, 106]), such that response times were longer on the relative pronoun when the preceding second noun was modified compared to when neither noun was modified. Again, since I predicted that an implicit prosodic boundary would be likely after the modified second noun, this potentially provides support for the idea that implicit prosodic boundaries can show up as reading slowdowns on the word after the location of the hypothesized boundary.

4.7.5 Discussion

Overall, this experiment replicated the main findings of Experiments 7 and 8, showing that they generalized to the configuration with the complex NP in object position rather than subject position. As in Experiment 7, modification of N2 led to more low attachment responses, and as in Experiment 8, modification of N1 led to more low attachment responses. Crucially, these effects also held true regardless of the length of the RC, which grants additional confidence in these findings by showing that these results obtain across multiple configurations. As discussed in Experiment 8, these results are inconsistent with the Modification Attraction Hypothesis, which states that interpretation responses should be driven by whichever noun was modified, and which would have predicted more high attachment responses in the Noun 1 Modification conditions. However, these results are consistent with the Visibility First Hypothesis, which states that attachment is primarily driven by prosodic visibility, and that modification can only influence attachment when two attachment sites are equally visible. This account can explain why Noun 1 Modification did not lead to more high attachment responses, because the first noun in this condition was also less visible than the second one, and so modification did not exert its influence.

The response times in this experiment also provide additional suggestive evidence that implicit prosodic boundaries might be reflected in a slowdown in response times, and that this slowdown occurs on the word *after* the position of the hypothesized boundary. In this experiment, I predicted that a modified Noun 1 would be more likely to be followed by an implicit prosodic boundary, and I found inflated response times on the following preposition in the Noun 1 Modification condition. I also predicted that a modified Noun 2 would be more likely to be followed by an implicit prosodic boundary, and I found inflated response times on the following relative pronoun in the Noun 2 Modification condition. However, I once again emphasize that further work is needed on the relationship between reading measures and implicit boundaries before we can confidently draw conclusions about implicit prosody from response times.

4.8 General Discussion

This chapter has investigated the assignment of implicit prosodic boundaries when either noun in the Complex NP + RC construction is modified. I first provided a detailed walkthrough of how an incremental parser would assign structure to the stimuli in Experiment 7, in which the second noun was modified. I then discussed the predictions about the effects of modification and boundaries on attachment made by different accounts, including the Repellent Boundaries Hypothesis, the Modification Attraction Hypothesis, and the Visibility First Hypothesis. Across three experiments, I found that modification of both the first noun and the second noun led to more high attachment interpretation responses. I argued that this was only compatible with the Visibility First Hypothesis, which states that prosodic visibility is the primary factor in the parser's attachment decisions, with a preference to attach to the most visible attachment site, and that modification and focus can only attract attachment when the two attachment sites are equally visible. While this is consistent with the results of the experiments, more work is needed to provide concrete support for this account. I now describe several potential follow-up studies.

In addition to the rapid prosodic transcription and production tasks proposed in the discussions of Experiments 7 and 8, a listening study would perhaps be the most straightforward way to hold boundaries constant - and therefore be certain of the prosodic structure on any given trial - and then test how the presence or absence of modification influences responses. For instance, a listening study could manipulate whether N1 or N2 is modified and whether a major prosodic break occurs after N1 or after N2, as in (110). If modification attracts attachment regardless of boundaries, then there should be an increase in N1 responses in (110b,d) compared to (110a,c). However, if visibility gates the use of modification, then the effect of modification should be reduced or nonexistent in (110c,d), in which N2 is more visible than N1, because attachment decisions should not be influenced by modification unless both potential attachment sites are equally visible.

(110)

a. The niece of the diligent waitress % who chided herself over the blunder...

b. The diligent niece of the waitress % who chided herself over the blunder...

c. The niece % of the diligent waitress who chided herself over the blunder...

d. The diligent niece % of the waitress who chided herself over the blunder...

The Visibility First Hypothesis should also generalize beyond modification. For instance, using the RC structure, one could use the same design as (110), but manipulate the presence or absence of an accent on N1 or N2 instead of modification,[4] or the presence or absence of a focus particle like *only*. If the effects of accents and focus particles are also gated by visibility in the same way as modification, then I would expect parallel results.

The Visibility First Hypothesis also makes predictions for structures beyond the complex NP + RC sequence. Recall that Carlson and Tyler (2018) tested the interaction of boundaries in various structures. However, they only considered cases where a boundary came after the second attachment site. One could replicate and extend their findings by testing what happens when a boundary intervenes between the two attachment sites. For instance, in the matrix vs. complement clause attachment ambiguity in (111), one could add conditions (111e) and (111f). If Carlson and Tyler (2018) are correct, and accents exert an influence regardless of boundaries, then there should be no interaction of boundary placement and accents: there should be a difference between (111e) and (111f), with more high attachment responses in (111e), because the role of the accent is not modulated by the boundary between the two attachment sites. However, the Visibility First Hypothesis would predict an interaction. Pitch accents would affect attachment preferences when there is a boundary after both attachment sites (111a,b) and when there is no boundary

[4]This would essentially be a replication of Lee and Garnsey's (2012) Experiment 2, discussed in Chapter 1; however, given Carlson and Potter's suggestion that this could be an accidental null, a replication would be prudent.

(111c,d) but not when the boundary intervenes between the two attachment sites (111e,f), because this boundary would render the first site less visible, and accents should only exert an influence when visibility does not distinguish potential attachment sites. Note that this would also be parallel to the interpretation I provided in Chapter 1 of Lee and Garnsey's (2012) findings, where the presence or absence of an accent on N2 in the complex NP + RC structure did not influence RC attachment preferences when an intonational phrase boundary intervened between N1 and N2.

(111) a. John CLAIMED that Mary arrived IPh last week.

b. John claimed that Mary ARRIVED IPh last week.

c. John CLAIMED that Mary arrived last week.

d. John claimed that MARY arrived last week.

e. John CLAIMED IPh that Mary arrived last week.

f. John claimed IPh that MARY arrived last week.

To summarize, the present findings only establish that the Repellent Boundary Hypothesis alone is insufficient (Experiments 7, 9) and that the Modification Attraction Hypothesis cannot account for the data without also acknowledging the role of prosodic boundaries (Experiments 8, 9). It is perhaps unsurprising that neither prosody nor information structure alone could explain the results of the experiments. However, the Visibility First Hypothesis is an important step forward because it makes an explicit claim about *how* the parser weighs these two sources of evidence. Testing its predictions will deepen our understanding of how information structure and prosody guide sentence processing.

Chapter 5

Conclusion

This dissertation has argued that research on implicit prosody must go beyond documenting that various prosodic features (e.g., stress, accents, boundaries) are represented in reading and turn to developing a theory of the incremental assignment of prosodic structure that specifies how and when various pressures like size, balance, and syntax-prosody correspondence exert their influence. This will require running more online reading studies. Eye-tracking while reading is the psycholinguistic task that most closely approximates natural reading situations, but collecting eye-tracking data often takes more time than running a self-paced reading or Maze study. Running eye-tracking studies is also not always an option, in particular in situations where the necessary equipment may not be available (e.g., running studies online or in the field), and was also not an option for this dissertation, because of the COVID-19 pandemic.

Thus, a major contribution of this dissertation is the demonstration that the Maze task is appropriate for studying implicit prosody: in Chapter 2, I replicated several major findings on both metrical and phrasal prosody. Prior to the replications in this dissertation, one could reasonably have doubted the Maze's ability to detect implicit prosodic effects, and/or expected that task-specific requirements would result in an implicit prosody that diverged from the prosody that readers assign in more natural reading situations. However, these replications show that this cannot be the case, because the same implicit prosodic effects are found in the Maze. In fact, we showed that the Maze is arguably preferable to self-paced reading even for implicit prosody, because the Maze and SPR comparisons showed many of the same advantages previously reported for the Maze, e.g., more localized effects that are not spread out over multiple spillover words and greater effect sizes (Boyce et al., 2020; Witzel et al., 2012).

The dissertation has also laid the groundwork for developing an incremental model of prosodic parsing. In Chapter 3, I provided an overview of the major grammatical constraints governing the syntax-prosody interface. Although grammatical theories of the interface typically take an entire sentence and consider all possible phrasings, the task of the parser is to incrementally build a structure word-by-word, without knowing what the final string will be. The question of incrementality has often been underexamined in this literature, and researchers have frequently proposed a default implicit prosody by considering which prosodic structure would be the best for the final string. I stressed the importance of considering the differences between first pass implicit prosody, assigned as the sentence unfolds one word at a time, and the final prosody, assigned once the parser has encountered all the material and has potentially revised the prosodic structure to achieve better syntax-prosody correspondence and balance.

Using a toy model of an incremental parser assigning structure to various sentences, I discussed how explicit models of prosodic parsing could translate the aforementioned grammatical constraints into principles that guide the parser's choices. For instance, BINMIN, which militates against phrases containing a single word, could be reflected in a principle we might call Prosodic Late Closure: a tendency to prefer adding incoming material into the current phonological phrase, in order to avoid a phrase that is too small. In contrast, BINMAX, which places a limit on how big phrases can be, could be implemented by making the probability of the parser inserting a prosodic break increase with each additional word that it places into the phrase currently being processed. The knowledge represented by constraints like EQUALSISTERS and STRONGSTART, which place requirements on the categories of sister constituents, could be used by the parser to make predictions about upcoming input. BALANCEDSISTERS could also be used predictively, such that the parser predicts the size of an upcoming phrase on the basis of the size of the phrase it just finished processing, but it might also be reflected in revision processes, since the parser may not be able to achieve balance until it has an idea of how many words are in the sentence, at which point it can change groupings as needed to ensure that words are distributed in a relatively even manner across phrases.

The experiments in Chapter 4 also provided a template for studying the incremental assignment of prosody: the important ingredients include (i) a task that can track online processing, such as the Maze, (ii) a manipulation that occurs prior to the change in implicit prosody that it induces, so that responses at or around the location of the hypothesized change in prosody can be analyzed to verify that this change is reflected in behavioral measures, and (iii) a dependent variable that should change if the implicit prosody manipulation would change attachment preferences. In Experiments 7 and 8, I examined whether the ambiguity advantage could be modulated with implicit prosody, reasoning that if a boundary between N2 and the RC led to more high attachment preferences, as predicted by the Repellent Boundary Hypothesis, then the high attachment penalty should be attenuated or reversed in this condition. However, because the length manipulations only ever *reinforced* English's pre-existing low attachment bias, we were not able to see any online impact of implicit prosody on the cost of high attachment: there was a high attachment penalty regardless of the length of the preceding complex NP. Still, the logic of the study shows the type of experimental design that could assess whether implicit prosody has affected first pass syntactic parsing.

The findings of Chapter 4 also suggested that implicit prosodic boundaries may be reflected in longer response times at the word immediately following the boundary. I contextualized this finding with respect to Staub (2007a), who suggested that readers may slow down on the word *after* a location where a prosodic boundary was expected but no comma appeared, indicating a disruption due to the violation of a prosodic expectation. However, I also described reasons why we might be skeptical about inferring the implicit prosodic structure based on response times. I proposed that future studies should incorporate a rapid prosodic transcription or production task, and then correlate the phrasings reported and/or produced by participants with the interpretation data and response times, in order to have more confidence in which phrasings are being assigned and how they relate to the different behavioral measures.

Finally, the fact that the hypothesized boundary between N2 and the RC in Experiments 7 and 9 did not lead to more high attachment preferences, but in fact resulted in the opposite, highlighted another important issue in this area of work: the manipulations that affect implicit prosody, such as the use of modification to increase constituent length and thereby induce a boundary, can also change other properties of the sentence like information structure. This complicates the task of inferring the implicit prosody from, e.g., end-of-sentence interpretations, because we must determine which changes in response are due to prosody and which are due to other factors. To this end, I proposed the Visibility First Hypothesis as a way of understanding how boundaries and accents could interact. This hypothesis states that attachment choices are primarily guided by prosodic visibility, with the parser preferring to attach incoming material to whichever attachment site is most prosodically visible (Schafer, 1997). However, when two attachment sites are equally visible, other sources of evidence, such as modification and focus, can affect attachment choices. I argued that the results of Experiments 7 through 9 were consistent with this hypothesis, while an account that predicted attachment responses entirely from the location of the hypothesized implicit boundaries, as in the Repellent Boundaries Hypothesis, or based solely on which noun was modified, as in the Modification Attraction Hypothesis, made incorrect predictions. I then proposed several follow-up studies to further test the Visibility First Hypothesis.

In closing, I would like to emphasize the importance of working towards a theory of incremental implicit prosodic assignment. Although many psycholinguists rely on data from reading studies, the prosodic component of processing is often overlooked.[1] Nevertheless, it is clear that many aspects of prosody are represented in silent reading, and that prosody can have consequences for core aspects of sen-

[1]This observation is not new, and is perhaps best conveyed by the title of Fodor (2002b): "Psycholinguistics cannot escape prosody." Still, it would serve us well to remind ourselves that, in the words of Taylor Swift, "You can hear it in the silence."

tence processing like attachment, reanalysis, and retrieval (Bader, 1998; Carlson et al., 2001; Cutler et al., 1997; Fodor, 1998, 2002a, 2002b; Frazier, Carlson, & Clifton Jr, 2006; Frazier et al., 2014; Kjelgaard & Speer, 1999; Royer, 2021; Wagner & Watson, 2010, i.a.). A theory that predicts the implicit prosody of novel sentences in advance would make it much easier for psycholinguists to consider how prosody could impact their results, and this dissertation has laid the groundwork for future theorizing and experimentation in this area.

Appendix A

Experiment 10: NP2 Modification in SPR

The interpretation data from Experiment 7 found that participants were more likely to provide low attachment interpretation responses when N2 was modified. To confirm that this result would replicate across tasks and was not Maze-specific, I ran a self-paced reading study using the same materials.

A.1 Methods

A.1.1 Materials

The same materials from Experiment 7 were used in Experiment 10.

A.1.2 Participants

54 native English speakers from the University of California, Santa Cruz participated in the experiment for course credit. Six participants were excluded because they failed to correctly respond to at least 70% of the comprehension questions following unambiguous sentences.

A.1.3 Procedure

The self-paced reading procedure described in previous experiments was also used in Experiment 10.

A.1.4 Results

A.1.4.1 Comprehension question responses

Table A.1 shows the percentage of high attachment interpretations in each condition. The overall pattern of results is similar to the results of Experiment 7, with the exception that participants tended to provide more incorrect responses in the unambiguous condition in the SPR experiment; this is perhaps unsurprising, because the Maze forces participants to more deeply process each sentence, whereas the SPR allows participants to proceed through the sentence by quickly tapping the space bar.

ATTACHMENT	LENGTH	% High Attachment
Ambiguous	Short	48.1 (4.2)
	Long	40.0 (3.4)
Low (N2)	Short	26.9 (3.0)
	Long	20.7 (2.6)
High (N1)	Short	68.3 (3.0)
	Long	68.1 (3.4)

Table A.1: % high attachment responses in Experiment 10

As in Experiment 7, I fit a Bayesian logistic mixed effects model to the response data with High Attachment, Ambiguity, Length, and the interactions of High Attachment and Length and of Ambiguity and Length as fixed effects, and the maximal random effects structure. There was a main effect of High Attachment (1.30, [1.00, 1.61]), with more high attachment interpretations in the High condition compared to Ambiguous and Low, a main effect of Ambiguity (-.60, [-.78, -.44]), with fewer high attachment interpretations in the Low condition compared to the Ambiguous condition, and a main effect of Length (-.15, [-.29, -.01]), with fewer high attachment responses when N2 was modified. No interactions were significant.

Thus, even though participants committed more errors in the present experiment, the qualitative pattern of results was quite similar to what was found in Experiment 7: the Long condition led to fewer high attachment interpretations, and responses in the unambiguous conditions were driven by the attachment height required by the reflexive. However, unlike Experiment 7, there was no interaction of High Attachment and Length in this study (.13, [-.05, .30]). Although the percentages suggest that the Length effect in this experiment was not present in the High condition, the lack of an interaction does not support this.

A.1.4.2 Response times

Observations shorter than 100ms or greater than 5000ms were excluded from the analysis, which was .3% of the data. As in Experiment 7, I collapsed across all Attachment conditions for analysis of the complex NP region. The mean RTs at each word in this region are shown in Figure A.1. Bayesian mixed effects models were fit to the response times at N2, the relative pronoun *who*, and the relative clause verb *chided*, with Length as the fixed effect and the maximal random effects structure. There was a main effect of Length at N2 (25.54, [7.70, 43.45]), replicating the finding from Experiment 7 that response times were slower at N2 in the Long condition. There was not a main effect of Length at the relative pronoun (0.93, [-15.85, 17.75]) nor at the relative clause verb (-6.40, [-21.63, 8.81]).

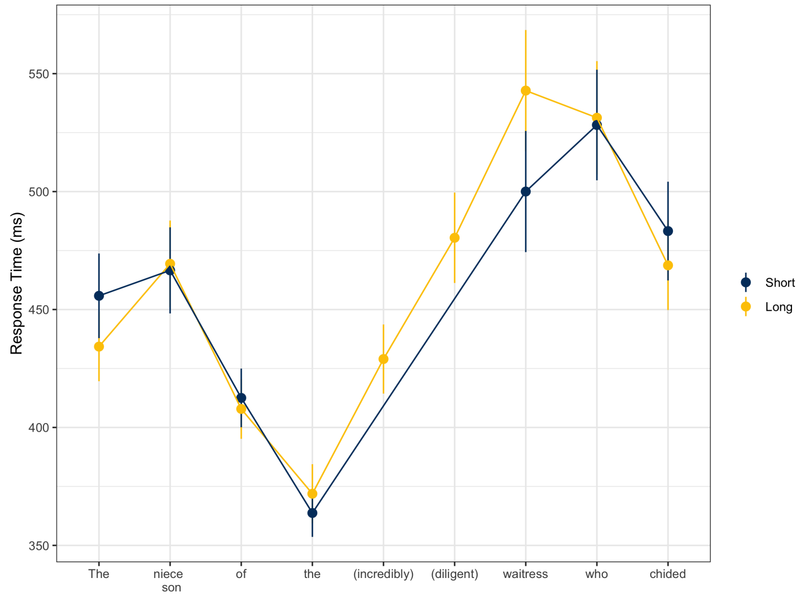


Figure A.1: Mean response times at each word in the complex NP in Experiment 10. Error bars show a 95% confidence interval.

Mean response times through the end of the relative clause are shown in Figure A.2. Again, I fit Bayesian mixed effects models to the response times at the reflexive (*themselves*) and each of the two spillover words (*over* and *the*) with High Attachment, Ambiguity, Length, and the interactions of Length and High Attachment and of Length and Ambiguity as fixed effects and the maximal random effects structure. There were no significant main effects or interactions at the reflexive. At the first spillover word, there was a main effect of High Attachment (46.77, [11.97, 81.57]), such that there was a slowdown in the High condition compared to the Low and Ambiguous conditions, and a main effect of Length (-17.06, [-31.47, -2.58]), such that response times were faster in the Long condition. At the second spillover word, there were similar effects of High Attachment (19.79, [3.36, 36.15]) and Length (-15.3, -28.34, -2.2]). No other main effects or interactions were significant.

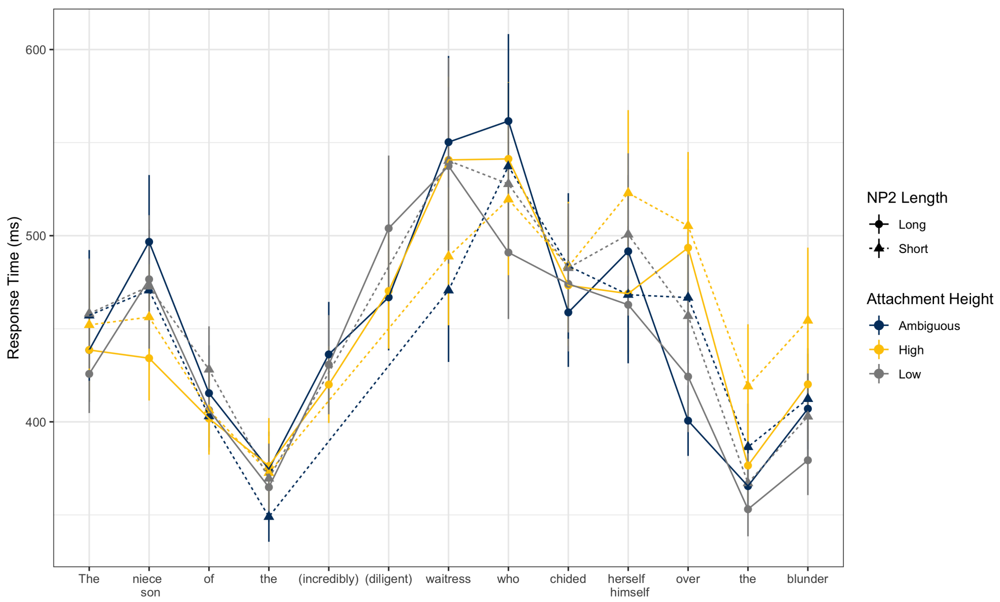


Figure A.2: Mean response times at each word through the relative clause in Experiment 10. Error bars show a 95% confidence interval.

A.1.5 Discussion

The findings from Experiment 10 largely replicated Experiment 7: there were more low attachment responses when N2 was modified, consistent with the Modification Attraction Hypothesis and the Visibility First Hypothesis. There was also a reading slowdown at N2 in the Long condition, potentially reflecting the hypothesized prosodic boundary in this position. However, I did not replicate the slowdown at the relative pronoun.

Appendix B

Experiment 1 and 2 Materials

Items (1) through (32) are from Breen and Clifton (2013), while items (33) through (40), as well as the foils for all items, were written for Experiments 1 and 2. The part of speech of the homograph (Noun, condition A, vs. Verb, condition B) was crossed with the homograph's stress pattern (Alternating vs. Non-alternating). For each condition, the Stress Mismatched foils are given in (i), the Mostly Stress Matched foils in (ii), and the Stress Matched foils in (iii).

(1) a. The brilliant {abstract/report} was accepted at the prestigious conference.

i. x-x-x persuade unfastens haven't grape isn't until entrust dishearten.

ii. x-x-x motivate {fastens/secures} have banana is from consider brighten.

iii. x-x-x motivate {fastens/secures} have banana is from consider brighten.

b. The brilliant {abstract/report} the best ideas from the things they read.

i. x-x-x persuade unfastens haven't shouldn't whoever kitten didn't unless banana myself.

ii. x-x-x motivate secures have should whom cat do and grape ours.

iii. x-x-x motivate secures have should whoever cat do and grape ours.

(2) a. The secretive {project/design} was closely guarded by the military

i. x-x-x establish awakens detect professor vinegar myself among therefore.

ii. x-x-x colonize {wakens/persuades} see teacher acids we from consequently.

iii. x-x-x colonize {wakens/persuades} see teacher acids we from consequently.

b. The secretive {project/design} an image of mystery and privacy.

i. x-x-x establish awakens detect prophet baffle intensify carry hence.

ii. x-x-x colonize persuades see saint err aggravate go furthermore.

iii. x-x-x colonize persuades see prophet err aggravate go furthermore.

(3) a. The vicious {combat/defeat} resulted in many casualties.

i. x-x-x inspect permeates disturb pumpkin multiply elevator.

ii. x-x-x hearten {moisten/pervades} unsettle gourd liven paragraph.

iii. x-x-x hearten {moisten/pervades} unsettle gourd liven paragraph.

b. The vicious {combat/defeat} their enemies with strength and malice.

i. x-x-x inspect permeates divide accumulate idiot saturate himself map.

ii. x-x-x hearten pervades halve elevate fool drench them atlas.

iii. x-x-x hearten pervades halve elevate fool drench them atlas.

(4) a. The unusual {permit/practice} was not sanctioned by the local government.

i. x-x-x analyze overfeeds devour carpet elephant herself into diminish hereby.

ii. x-x-x infatuate fattens eat rug zebra me to worsen furthermore.

iii. x-x-x infatuate fattens eat rug zebra me to worsen furthermore.

b. The unusual {permit/practice} the behaviors that most people find unacceptable.

i. x-x-x analyze overfeeds devour rationalize forget doesn’t enough yourself elevators.

ii. x-x-x infatuate {fattens/expands} eat justify lose does many yours bibliography.

iii. x-x-x infatuate {fattens/expands} eat establish lose does many yours bibliography.

(5) a. The strong {contrast/partner} was hard to ignore.

i. x-x-x exist illuminates unsee simulate could’ve seventy.

ii. x-x-x were brightens see feign can’t thirteen.

iii. x-x-x were brightens see feign can’t thirteen.

b. The strong {contrast/partner} with their weaker friends.

i. x-x-x exist illuminates unsee simulate agonize never.

ii. x-x-x were {adorns/brightens} see feign suffer not.

iii. x-x-x were {adorns/brightens} see feign suffer not.

(6) a. The handy {object/answer} was helpful to the student.

i. x-x-x emit disables whittle century herself after each.

ii. x-x-x lighten injures carve decade me for every.

iii. x-x-x lighten injures carve decade me for every.

b. The handy {object/answer} to their less able superiors.

i. x-x-x emit disables whittle along herself disorganize complicate.

ii. x-x-x lighten {impedes/injures} carve from me scramble embarrass.

iii. x-x-x lighten {impedes/injures} carve from me scramble embarrass.

(7) a. The awkward {subject/challenge} was discussed by the family.

i. x-x-x examine discovers undo prosecute robin among occur.

ii. x-x-x blacken locates don't accuse crow with happen.

iii. x-x-x blacken locates don't accuse crow with happen.

b. The awkward {subject/challenge} themselves to difficult tasks.

i. x-x-x examine discovers authorize invoke regardless among

ii. x-x-x blacken detects endorse cite anyhow with.

iii. x-x-x blacken detects endorse cite anyhow with.

(8) a. The phony {address/return} was discovered by the authorities.

i. x-x-x portray wherever defy accumulation whoever into reinforce.

ii. x-x-x sadden {whether/unless} go collection who in consolidate.

iii. x-x-x sadden {whether/unless} go collection who in consolidate.

b. The phony {address/return} the messages with fake sincerity.

i. x-x-x portray wherever defy illuminate whoever into reinforce.

ii. x-x-x sadden unless go clarify who in consolidate.

iii. x-x-x sadden unless go clarify who in consolidate.

(9) a. The epic {details/reports} of Greek heroes are well known.

i. x-x-x impair wherever isn't Aspire contribute piglet shouldn't bargains.

ii. x-x-x suffer {whether/unless} is Yearn donate pig should costs.

iii. x-x-x suffer {whether/unless} is Yearn donate pig should costs.

b. The epic {details/reports} the story of Odysseus' journey.

i. x-x-x impair wherever isn't contribute myself fantasizes meanwhile.

ii. x-x-x suffer unless is donate me catastrophizes during.

iii. x-x-x suffer unless is donate me catastrophizes during.

(10) a. The local {contests/questions} occupied the entire street.

i. x-x-x assemble whenever affirms deprive subsidize neither.

ii. x-x-x whiten whether testifies lose endow both.

iii. x-x-x whiten whether testifies lose endow both.

b. The local {contests/questions} the new zoning laws.

i. x-x-x assemble whenever hasn't deprive tolerate neither.

ii. x-x-x whiten {unless/whether} has lose suffer both.

iii. x-x-x whiten {unless/whether} has lose suffer both.

(11) a. The subtle {conduct/disguise} of the secret agent alerted no one to his presence.

i. x-x-x relies develops haven't decide improve condense horrify carafe wouldn't myself along both.

ii. x-x-x dampens {widens/extends} have opt sharpen thicken unsettle jug would me with neither.

iii. x-x-x dampens {widens/extends} have opt sharpen thicken unsettle jug would me with neither.

b. The subtle {conduct/disguise} themselves in a manner that does not look suspicious.

i. x-x-x relies develops attest decide obtain condense lady carafe wouldn't myself receptacle.

ii. x-x-x dampens extends vouch opt get thicken lord jug would me container.

iii. x-x-x dampens extends attest opt get thicken lord jug would me container.

(12) a. The hefty {decrease/budget} was a surprise to the school board.

i. x-x-x restore oppresses undo deny establish ignore without couldn't within.

ii. x-x-x freshen darkens do err entrust tempt with could with.

iii. x-x-x freshen darkens do err entrust tempt with could with.

b. The hefty {decrease/budget} their calories in an effort to lose weight.

i. x-x-x restore oppresses undo infuriate coerce ignore throughout couldn't within sandal.

ii. x-x-x freshen {conceals/darkens} do petrify doze tempt during could with shoe.

iii. x-x-x freshen {conceals/darkens} do petrify doze tempt during could with shoe.

(13) a. The intelligent convict exhibited surprising coherence.

i. x-x-x provoke solidifies fantasize reconcile furthermore.

ii. x-x-x antagonize hardens preoccupy resettle moreover.

iii. x-x-x antagonize hardens preoccupy resettle moreover.

b. The intelligent {convict/suspect} the criminals after a fair trial.

i. x-x-x provoke solidifies annoy administer enlist gotten into not.

ii. x-x-x antagonize {congeals/hardens} go regulate whiten been on never.

iii. x-x-x antagonize {congeals/hardens} go regulate whiten been on never.

(14) a. The witty {addict/puzzle} was a hit at all of the parties.

   i. x-x-x bestow elongate recite isn't hasn't exist into avoid neither fiend.

   ii. x-x-x donate lengthen cite is has be to lose nor demon.

   iii. x-x-x donate lengthen cite is has be to lose nor demon.

   b. The witty {addict/puzzle} their friends with their jokes and stories.

   i. x-x-x bestow elongate recite isn't hasn't exist into avoid furthermore.

   ii. x-x-x donate {extends/lengthen} cite is has be to lose therefore.

   iii. x-x-x donate {extends/lengthen} cite is has be to lose therefore.

(15) a. The lovely {entrance/delight} was appreciated by all of the guests.

   i. x-x-x afford depresses incite editorializes isn't into puppy haven't delete.

   ii. x-x-x moisten {saddens/destroys} cite decolonizes is to dog have flout.

   iii. x-x-x moisten {saddens/destroys} cite decolonizes is to dog have flout.

   b. The lovely {entrance/delight} the people who see them arrive.

   i. x-x-x afford depresses incite emancipate isn't into puppy eleven.

ii. x-x-x moisten destroys cite sharpen is to dog fifteen.

iii. x-x-x moisten destroys cite sharpen is to dog fifteen.

(16) a. The speedy {relay/process} was finished very quickly.

i. x-x-x subdue solidifies inquire giraffe encourage zucchini.

ii. x-x-x chasten stiffens ask zebra hearten carrot.

iii. x-x-x chasten stiffens ask zebra hearten carrot.

b. The speedy {relay/process} the news as soon as it arrives.

i. x-x-x subdue solidifies inquire yourself kitten haven't yourself jersey prior.

ii. x-x-x chasten {promotes/stiffens} ask they cat have you shirt before.

iii. x-x-x chasten {promotes/stiffens} ask they cat have you shirt before.

(17) a. The mysterious {invite/notice} was the topic of much gossip.

i. x-x-x dominate reduces forget forehead humidify powder unto eternally.

ii. x-x-x monopolize deadens lose head moisten dust by always.

iii. x-x-x monopolize deadens lose head moisten dust by always.

b. The mysterious {invite/notice} the scrutiny of others.

i. x-x-x dominate reduces forget evaluate defy eternally.

ii. x-x-x monopolize {absorbs/deadens} lose analyze flout always.

iii. x-x-x monopolize {absorbs/deadens} lose analyze flout always.

(18) a. The foreign {import/visit} was a hot topic of conversation around town.

i. x-x-x demolish consolidates retain forget decides before bestow emancipate inherit between.

ii. x-x-x flatten tightens lose bathe opts after give disenfranchise receive for.

iii. x-x-x flatten tightens lose bathe opts after give disenfranchise receive for.

b. The foreign {import/visit} the things they miss from home.

i. x-x-x demolish consolidates retain forget decides into bestow decade.

ii. x-x-x flatten {secures/tightens} lose bathe opts to give year.

iii. x-x-x flatten {secures/tightens} lose bathe opts to give year.

(19) a. The vulgar {insult/surprise} was a shock to the audience.

i. x-x-x expire encourages adopt couldn't chooses evade without regardless.

ii. x-x-x toughen {heartens/elates} choose could opts lose with anyhow.

iii. x-x-x toughen {heartens/elates} choose could opts lose with anyhow.

b. The vulgar {insult/surprise} the people with their inappropriate comments.

i. x-x-x expire encourages adopt aggravate chooses evade redecorate furthermore.

ii. x-x-x toughen elates choose worsen opts lose overcompensate therefore.

iii. x-x-x toughen elates choose worsen opts lose overcompensate therefore.

(20) a. The angry {protest/lament} was closely watched by the police.

i. x-x-x pursue diminishes appoint observe simulate haven't below understand.

ii. x-x-x follow {weakens/depletes} choose brighten feign have of believe.

iii. x-x-x follow {weakens/depletes} choose brighten feign have of believe.

b. The angry {protest/lament} their treatment by the police and courts.

i. x-x-x pursue diminishes appoint observe simulate haven't examine myself whenever.

ii. x-x-x follow depletes choose brighten feign have observe me when.

iii. x-x-x follow depletes choose brighten feign have observe me when.

(21) a. The exotic {extract/bottle} was sold at the expensive shop.

i. x-x-x ignite alleviates decide could've isn't unto obliterate whichever.

ii. x-x-x rekindle softens opt could is for demolish which.

iii. x-x-x rekindle softens opt could is for demolish which.

b. The exotic {extract/bottle} the flavors of herbs and flowers for cooking.

i. x-x-x ignite alleviates decide diminish isn't unto donates beside acquires above.

ii. x-x-x rekindle {relieves/softens} opt deaden is for gives versus gets under.

iii. x-x-x rekindle {relieves/softens} opt deaden is for gives versus gets under.

(22) a. The righteous {rebel/struggle} was supported by a foreign government.

i. x-x-x petrify detaches refer renovate require hasn't adjust regardless.

ii. x-x-x frighten loosens cite refurbish want has loosen anyway.

iii. x-x-x frighten loosens cite refurbish want has loosen anyway.

b. The righteous {rebel/struggle} against oppressive regimes and dictators.

i. x-x-x petrify detaches delineate renovate distribute hasn't regardless.

ii. x-x-x frighten {unhooks/loosens} pertain refurbish dispense has anyway.

iii. x-x-x frighten {unhooks/loosens} pertain refurbish dispense has anyway.

(23) a. The sneaky {pervert/escape} was featured on all the local news channels.

i. x-x-x condense revitalizes conclude forgive covert defy myself each receive ago.

ii. x-x-x shorten {freshens/restores} opt deepen yearn err them every earn during.

iii. x-x-x shorten {freshens/restores} opt deepen yearn err them every earn during.

b. The sneaky {pervert/escape} the truth when it stands in their way.

i. x-x-x condense revitalizes conclude unsee covet defy myself doesn't receive hasn't.

ii. x-x-x shorten restores opt see yearn err them does earn has.

iii. x-x-x shorten restores opt see yearn err them does earn has.

(24) a. The unfortunate {reject/recruit} was run out of town.

i. x-x-x humanize illuminates moisten hasn't kitten misplace mallard.

ii. x-x-x personify {lightens/expounds} bathe has cat lose duck.

iii. x-x-x personify {lightens/expounds} bathe has cat lose duck.

b. The unfortunate {reject/recruit} the offers of help from strangers.

i. x-x-x humanize illuminates moisten purify kitten misplace mallard width.

ii. x-x-x personify expounds bathe sweeten cat lose duck thickness.

iii. x-x-x personify expounds bathe sweeten cat lose duck thickness.

(25) a. The evil suspect was distressing to the public.

i. x-x-x expand wherever disinfect fertilize heifer atop exceed.

ii. x-x-x ripen whether bathe recycle cow at strengthen.

iii. x-x-x ripen whether bathe recycle cow at strengthen.

b. The evil suspect their victims will be freed by the police.

i. x-x-x expand wherever disinfect condense heifer atop hasn't ascend into seem.

ii. x-x-x ripen {unless/whether} bathe shorten cow at has soar to appear.

iii. x-x-x ripen {unless/whether} bathe shorten cow at has soar to appear.

(26) a. The sophisticated {survey/design} was used by many researchers.

i. x-x-x theorizes fortifies persuade should've isn't awhile wherefore.

ii. x-x-x philosophizes {coarsens/congeals} coax should is during furthermore.

iii. x-x-x philosophizes {coarsens/congeals} coax should is during furthermore.

b. The sophisticated {survey/design} their surroundings with a critical eye.

i. x-x-x theorizes fortifies persuade overshadow isn't should've wherefore onto.

ii. x-x-x philosophizes congeals coax outnumber is should furthermore on.

iii. x-x-x philosophizes congeals coax outnumber is should furthermore on.

(27) a. The religious {convert/reform} was applauded by many citizens.

i. x-x-x restock contaminates convince unload haven't in demonstrate.

ii. x-x-x replenish {blackens/pollutes} coax deposit have into testify.

iii. x-x-x replenish {blackens/pollutes} coax deposit have into testify.

b. The religious {convert/reform} the beliefs of their friends and neighbors.

i. x-x-x restock contaminates convince give haven't undo perform warthog variation.

ii. x-x-x replenish pollutes coax bestow have drench sing hog difference.

iii. x-x-x replenish pollutes coax bestow have drench sing hog difference.

(28) a. The weekly {inserts/offers} were particularly annoying this month.

i. x-x-x detach anyhow withers reduce fairy giraffe aware.

ii. x-x-x loosen unless weakens lessen elf zebra mindful.

iii. x-x-x loosen unless weakens lessen elf zebra mindful.

b. The weekly {inserts/offers} useless coupons in every issue.

i. x-x-x detach anyhow withers reduce fairy giraffe aware.

ii. x-x-x loosen {unless/whether} weakens lessen elf zebra mindful.

iii. x-x-x loosen {unless/whether} weakens lessen elf zebra mindful.

(29) a. The sturdy {compress/bandage} was used to stop the bleeding.

i. x-x-x attach purifies entice decide isn't besides ascend nor.

ii. x-x-x fasten whitens tempt opt is else soar neither.

iii. x-x-x fasten whitens tempt opt is else soar neither.

b. The sturdy {compress/bandage} the wounds of their fallen comrades.

i. x-x-x attach purifies entice decide isn't besides while nor.

ii. x-x-x fasten {refines/whitens} tempt opt is else during neither.

iii. x-x-x fasten {refines/whitens} tempt opt is else during neither.

(30) a. The suspicious {record/shadow} was discovered to be fake.

i. x-x-x organize permeates attract monkey isn't within every.

ii. x-x-x unscramble dampens tempt gorilla is with each.

iii. x-x-x unscramble dampens tempt gorilla is with each.

b. The suspicious {record/shadow} their friends and colleagues.

i. x-x-x organize permeates attract persuade isn't regularly.

ii. x-x-x unscramble {pervades/dampens} tempt tempt is always.

iii. x-x-x unscramble {pervades/dampens} tempt coax is always.

(31) a. The tacky {present/display} was covered in sequins and feathers.

i. x-x-x enlarge enlarges sharpen solidify isn't expand carafe hence.

ii. x-x-x broaden {strengthens/expands} whet stiffen is widen jug therefore.

iii. x-x-x broaden {strengthens/expands} whet stiffen is widen jug therefore.

b. The tacky {present/display} themselves tastelessly when in public.

i. x-x-x enlarge enlarges animate frighten isn't expand hence.

ii. x-x-x broaden expands believe terrify is widen therefore.

iii. x-x-x broaden expands believe terrify is grow therefore.

(32) a. The modest {discount/value} was appreciated by the customers at the store.

i. x-x-x alleviate demolishes jostle legalize haven't unto adore doggy along exist.

ii. x-x-x sweeten flattens cram decriminalize have from idolize pup for be.

iii. x-x-x sweeten flattens cram decriminalize have from idolize pup for be.

b. The modest {discount/value} the importance of their ideas and opinions.

   i. x-x-x alleviate demolishes jostle confuse haven't unto commend doggy besides.

   ii. x-x-x sweeten {destroys/flattens} cram befuddle have from admire pup moreover.

   iii. x-x-x sweeten {destroys/flattens} cram befuddle have from admire pup moreover.

(33) a. The powerful {protest/debate} was covered on the news.

   i. x-x-x administer enhances collide agonize hasn't unto neither.

   ii. x-x-x tranquilize {deepens/promotes} cram suffer has for nor.

   iii. x-x-x tranquilize {deepens/promotes} cram suffer has for nor.

b. The powerful {protest/debate} the policies of the opposition.

   i. x-x-x administer enhances collide inflame hasn't unto overcompensate.

   ii. x-x-x tranquilize promotes cram aggravate has for misinterpret.

   iii. x-x-x tranquilize promotes cram aggravate has for misinterpret.

(34) a. The corrupt {pervert/attack} was exposed by the journalist.

i. x-x-x cherish contributes descend purify unsee yours shouldn't each.

ii. x-x-x enjoy {donates/endows} deign ordain see in elevate.

iii. x-x-x enjoy {donates/endows} deign ordain see in elevate.

b. The corrupt {pervert/attack} the system to protect their interests.

i. x-x-x cherish contributes descend purify unsee yours shouldn't each.

ii. x-x-x enjoy endows deign whiten see yourself should every.

iii. x-x-x enjoy endows deign whiten see yourself should every.

(35) a. The generous {escort/usher} was appreciated by the patrons.

i. x-x-x appease elongates unwind nauseate isn't under firstly.

ii. x-x-x pacify lengthens doze sensationalize is on primarily.

iii. x-x-x pacify lengthens doze sensationalize is on firstly.

b. The generous {escort/usher} the patrons to their seats.

i. x-x-x appease elongates unwind offend isn't under fifty.

ii. x-x-x pacify {extends/lengthens} doze brighten is on five.

iii. x-x-x pacify {extends/lengthens} doze brighten is on five.

(36) a. The helpful {impact/support} of generous donors was appreciated.

i. x-x-x revere diminishes abort distribute fortify doggy emerge.

ii. x-x-x cherish {weakens/depletes} fail allocate toughen pup materialize.

iii. x-x-x cherish {weakens/depletes} fail allocate toughen pup materialize.

b. The helpful {impact/support} their community through volunteer work.

i. x-x-x revere diminishes abort speculate doggy compose either.

ii. x-x-x cherish depletes fail hypothesize pup decompose both.

iii. x-x-x cherish depletes fail hypothesize pup decompose both.

(37) a. The careful {recall/review} had identified contaminated onions.

i. x-x-x inflame discovers undo execute dishearten have.

ii. x-x-x worsen {locates/detects} err assassinate demoralize haven't.

iii. x-x-x worsen {locates/detects} err assassinate demoralizes haven't.

b. The careful {recall/review} the important information for exams.

i. x-x-x inflame discovers undo anyway examine heifer communicate.

ii. x-x-x worsen detects err moreover reconsider cow impart.

iii. x-x-x worsen detects err moreover reconsider cow impart.

(38) a. The insensitive {insult/command} was condemned by the executive.

i. x-x-x decompose develops imitate publicize isn't onto organize.

ii. x-x-x disintegrate {widens/extends} feign announce is to reorganize.

iii. x-x-x disintegrate {widens/extends} feign announce is to reorganize.

b. The insensitive {insult/command} their subordinates when managing projects.

i. x-x-x decompose develops imitate reconstruct isn't giraffe banana.

ii. x-x-x disintegrate extends feign rejuvenate is elephant apple.

iii. x-x-x disintegrate extends feign rejuvenate is elephant apple.

(39) a. The irresponsible {increase/neglect} was criticized by the accountants.

i. x-x-x paralyze awakens submerge appall myself along impair.

ii. x-x-x incapacitate {wakens/persuades} drench horrify me from disable.

iii. x-x-x incapacitate {wakens/persuades} drench horrify me from disable.

b. The irresponsible {increase/neglect} the budget and ignore the consequences.

i. x-x-x paralyze awakens submerge appall myself zebra wouldn't creepily.

ii. x-x-x incapacitate persuades drench frighten me giraffe would threateningly.

iii. x-x-x incapacitate persuades drench frighten me giraffe would threateningly.

(40) a. The hopeless {reject/request} was ignored by everyone involved.

i. x-x-x condense unfastens defy acknowledge yourself however driver.

ii. x-x-x thicken {fastens/secures} flout confess them anyway chauffeur.

iii. x-x-x thicken {fastens/secures} flout confess them anyhow chauffeur.

b. The hopeless {reject/request} the guidance of their peers.

i. x-x-x condense unfastens defy each yourself along every.

ii. x-x-x thicken secures flout every them from the.

iii. x-x-x thicken secures flout every them from the.

Appendix C

Experiment 3 and 4 Materials

As described in Chapter 3, the items for Experiment 3 and 4 were adapted from previous work on the NP/Z garden path (Anderson & Carlson, 2010; Christianson et al., 2001; Frazier, Carminati, et al., 2006; Frazier & Rayner, 1982; Staub, 2007a, 2007b).

The structure of the sentence (NP vs. Z) was crossed with the presence of punctuation between the subordinate clause and the matrix clause (Comma vs. No Comma). For each condition, the Stress Mismatched foils are given in (i), the Mostly Stress Matched foils in (ii), and the Stress Matched foils in (iii).

(1) As John hunted(,) the frightened rabbit(,) (the turkey) escaped through the woods.

   i. x-x-x Table computer undo however what(,) (forget afford) deliver cupcake lady ergo.

ii. x-x-x Book racket(,) have therefore whither(,) (lose tighten) portray cake guy hence.

iii. x-x-x Book racket(,) have therefore whither(,) (lose tighten) portray cake guy hence.

(2) After Anne visited(,) the British relatives(,) (the cousins) moved to the countryside.

i. x-x-x Portray protein(,) bottle Anymore inclusive(,) (doesn't) those veggie defy baby friendly.

ii. x-x-x Weep algebra(,) bag Whether troublesome(,) (does any) rug pry sir heavenly.

iii. x-x-x Weep algebra(,) bag Whether troublesome(,) (does any) rug pry sir heavenly.

(3) Even when Todd cleaned(,) the small kitchen(,) (the room) smelled like old garbage.

i. x-x-x handle Jacket meadow(,) ocean couldn't explores(,) (waddle until) wallop pencil gotten confess.

ii. x-x-x door Vest hill(,) sky could broadens(,) (jump from) slap tee been coarsen.

iii. x-x-x door Vest hill(,) sky could broadens(,) (jump from) slap tee been coarsen.

(4) While the clown juggled(,) the sharp knives(,) (the apples) remained on the table.

i. x-x-x undo avow askew(,) candle conveys revise(,) (kitten from) kayak taxi couldn't eight.

ii. x-x-x eat have glottal(,) rug grows hear(,) (cat during) canoe bus could seven.

iii. x-x-x eat have glottal(,) rug grows hear(,) (cat during) canoe bus could seven.

(5) While Thomas walked(,) the anxious poodle(,) (the collie) barked at the neighbors.

i. x-x-x Barista virgin(,) cushion however encode(,) (donut ordain) repeat undo unsee communism.

ii. x-x-x Seamstress berth(,) fund whereby worsen(,) (sale blacken) say do see kindness.

iii. x-x-x Seamstress berth(,) fund under worsen(,) (sale blacken) say do see kindness.

(6) As Janet baked(,) the doughy bread(,) (the brownies) cooled on a rack.

i. x-x-x Tree shoulder(,) atlas suggest veranda(,) (rearm defend) journey myself unrip resets.

ii. x-x-x Typo purse(,) map redden grass(,) (lose blacken) trip we rip sets.

iii. x-x-x Typo purse(,) map redden grass(,) (lose blacken) trip we rip sets.

(7) As the kid drank(,) the sweet juice(,) (the milk) dripped on the counter.

i. x-x-x refill shouldn't karma(,) unit exist resound(,) (refund deny) bellow haven't wouldn't anymore.

ii. x-x-x fill should faith(,) link breathe solve(,) (fund nab) scream have would neither.

iii. x-x-x fill should faith(,) link breathe solve(,) (fund nab) scream have would neither.

(8) While the violinist practiced(,) the difficult concerto(,) (the symphony) blared from the radio.

i. x-x-x lady continues financier(,) liver deactivate emphasize(,) (woman indulge) aviary okay title whichever.

ii. x-x-x guy discontinue grocer(,) nose activate imagine(,) (man tolerate) coop bad name which.

iii. x-x-x guy discontinue grocer(,) nose activate imagine(,) (man tolerate) coop bad name anyway.

(9) While the thief hid(,) the elegant jewelry(,) (a diamond) sparkled in the light.

i. x-x-x buddy couldn't apple(,) decade however avoid(,) (myself calzones) volcano piglet undo cannot.

ii. x-x-x pal could plum(,) year anyway argue(,) (me pizzas) geyser pig do can't.

iii. x-x-x pal could plum(,) year anyway argue(,) (me pizzas) geyser pig do can't.

(10) As Michael smoked(,) the expensive cigar(,) (the pipe) glowed in the dark.

i. x-x-x Deserve oceans(,) atlas anyhow widen(,) (decade forgot) rodeo puppy liver defy.

ii. x-x-x Nourish lakes(,) map regardless enjoy(,) (year sang) gig dog heart pry.

iii. x-x-x Nourish lakes(,) map regardless enjoy(,) (year sang) gig dog heart pry.

(11) While the skipper sailed(,) the old ship(,) (the dinghy) leaked out the back.

i. x-x-x caress inquire demon(,) ankle ago himself(,) (bowtie explore) govern deity unsee wherever.

ii. x-x-x kiss widen fiend(,) heel since them(,) (sock freshen) bawl god see where.

iii. x-x-x kiss widen fiend(,) heel since them(,) (sock freshen) bawl god see where.

(12) Whenever the instructor taught(,) the lazy students(,) (the class) ignored the lecture.

i. x-x-x rodent escalate luxury(,) dismiss devour what(,) (yourself exist) elephant raisin eleven.

ii. x-x-x mouse determine wealth(,) miss fatten whether(,) (us breathe) giraffe peach twenty.

iii. x-x-x mouse determine wealth(,) miss fatten whether(,) (us breathe) giraffe peach twenty.

(13) Because the baby clutched(,) his loving mother(,) (the family) stayed until the end.

i. x-x-x bottle whomever vanquish(,) rodent subdue steam(,) (swindle destroy) freshen onset across neither.

ii. x-x-x mug whom trounce(,) mouse suffer vapor(,) (hoax freshen) wash begin from nor.

iii. x-x-x mug someone trounce(,) mouse suffer vapor(,) (hoax freshen) wash begin from nor.

(14) Because the dog ate(,) the helpful medicine(,) (the treatment) had its effect.

i. x-x-x lady whomever season(,) target predict moreover(,) (infant alongside) divide unsee emerges.

ii. x-x-x lord whom span(,) goal suffer furthermore(,) (kid during) fork see begins.

iii. x-x-x lord whom span(,) goal suffer furthermore(,) (kid during) fork see begins.

(15) After the lifeguard lectured(,) the eager swimmers(,) (the group) jumped into the pool.

i. x-x-x warthog should endanger(,) bottle embark however(,) (vessel whether) chipmunk lord whether sudden.

ii. x-x-x hog shouldn't donate(,) flask redden therefore(,) (vase when) squirrel lady if fast.

iii. x-x-x hog shouldn't donate(,) flask redden therefore(,) (vase when) squirrel lady if fast.

(16) While the woman cooked(,) the delicious pasta(,) (the pot) boiled on the stove.

i. x-x-x engage memorize autumn(,) gather recognize how(,) (refund unto) otter unit peruse seven.

ii. x-x-x lose realize spring(,) met deforest whether(,) (fund from) bear link browse six.

iii. x-x-x lose realize spring(,) met deforest whether(,) (fund from) bear link browse six.

(17) As the gangster shot(,) the opposing gang(,) (the boss) fell to the ground.

i. x-x-x dollar mitigate listen(,) vision where manage(,) (carrot among) heifer flower kitten either.

ii. x-x-x cent lessen hear(,) dream wherever cope(,) (squash for) cow plant cat and.

iii. x-x-x cent lessen hear(,) dream wherever cope(,) (squash for) cow plant cat and.

(18) While Mary was mending(,) the grandfather clock(,) (the doorbell) chimed in the hall.

i. x-x-x Basin duty became(,) dribble wherever haven't(,) (lose solvents) eagle taxi undo receive.

ii. x-x-x Sink task drank(,) drop anywhere have(,) (lose solvents) hawk car do earn.

iii. x-x-x Basin task strengthen(,) drop anywhere have(,) (lose carry) hawk car do earn.

(19) While the audience cheered(,) the flamboyant magician(,) (his partner) waved a baton.

i. x-x-x pitcher moreover harmonica(,) burrow presuppose enshrines(,) (erase endow) harness table freeze.

ii. x-x-x jug anyway lute(,) nest awakens considers(,) (kill frighten) leash box engorge.

iii. x-x-x jug anyway lute(,) nest awakens considers(,) (kill frighten) leash box engorge.

(20) After the patient asked(,) the caring nurse(,) (the doctor) shared the news.

i. x-x-x river have cactus(,) duty invest waffle(,) (defy distill) beetle bowtie doesn't.

ii. x-x-x pool haven't sand(,) quest brighten egg(,) (lose shorten) ant hat don't.

iii. x-x-x pool haven't sand(,) quest brighten egg(,) (lose shorten) ant hat don't.

(21) Before the customer interrupted(,) the busy manager(,) (the cashier) completed the transaction.

i. x-x-x couldn't moreover dancer(,) listen give deserve(,) (wallop inhabit) entrust kitten undoubtedly.

ii. x-x-x could anywhere ballerina(,) hear bestow tenderize(,) (slap settle) deliver cat completely.

iii. x-x-x could anywhere ballerina(,) hear soften tenderize(,) (slap evolve) deliver cat completely.

(22) As long as the king governed(,) his many subjects(,) his advisors feared his wrath.

i. x-x-x collar undo amount whoever motherland(,) arise confess document(,) (defies unsee) loser unsee many.

ii. x-x-x leash do sum whom homeland(,) rise whiten download(,) (eats unburden) dork see much.

iii. x-x-x leash do sum whom homeland(,) rise whiten download(,) (eats unburden) dork see much.

(23) As the guard investigated(,) the sneaky thief(,) (the accomplice) grabbed the money.

i. x-x-x peanut entrust quantitative(,) girdle provide carrot(,) (abhor devastate) aspire skillet shouldn't.

ii. x-x-x nut give degenerative(,) belt whiten pea(,) (loathe mismanage) wish pan should.

iii. x-x-x nut give degenerative(,) belt whiten pea(,) (loathe mismanage) wish pan shouldn't.

(24) When the dog scratched(,) the irritated vet(,) (her assistant) removed the muzzle.

i. x-x-x cabin neither hindsight(,) liver emphasize gumbo(,) (flicker confuse) optimize statue forever.

ii. x-x-x house nor stance(,) lung cannibalize soup(,) (flash befuddle) improve vase never.

iii. x-x-x house nor stance(,) lung cannibalize soup(,) (flash befuddle) improve vase never.

(25) While Eric played(,) the grand piano(,) (the harp) snapped a string.

i. x-x-x Inch minute(,) gotten couldn't hydrate(,) (snivel myself) listen river beyond.

ii. x-x-x Meter year(,) been could dehydrate(,) (weep our) hear sea with.

iii. x-x-x Meter year(,) been could dehydrate(,) (weep our) hear sea with.

(26) As the principal lectured(,) the graduating seniors(,) (the parents) planned a party.

i. x-x-x award examine disprove(,) sunup reciprocate dumb(,) (heifer hence) hasn't deny old.

ii. x-x-x gift irritate donate(,) dawn radicalize stupid(,) (cow therefore) has say mature.

iii. x-x-x gift irritate donate(,) dawn radicalize stupid(,) (cow therefore) has say mellow.

(27) While Pamela sketched(,) the little girl(,) (the dog) bothered the cat.

i. x-x-x Offend garbage(,) undo unless afford(,) (carafe couldn't) prohibit ballot despite.

ii. x-x-x Emphasize trash(,) do after hear(,) (jug could) hamper vote but.

iii. x-x-x Emphasize trash(,) do after hear(,) (jug could) hamper vote but.

(28) After the soldiers saluted(,) the new general(,) (the officer) ended the exercise.

i. x-x-x wander situate berry(,) defrost pursue imagine(,) (unsee remember) yearn shower therefore.

ii. x-x-x hike strengthen banana(,) melt seek maximize(,) (see memorize) covet rain furthermore.

iii. x-x-x hike strengthen banana(,) melt seek maximize(,) (see memorize) covet rain furthermore.

(29) As the class watched(,) the physics teacher(,) (a student) asked a question.

i. x-x-x idea along helpful(,) doesn't justify believe(,) (myself hence) wherefore fabric either.

ii. x-x-x plan from sweet(,) does soften donate(,) (me therefore) how cloth both.

iii. x-x-x plan from sweet(,) does soften donate(,) (me therefore) how cloth either.

(30) While the chimps groomed(,) the hairy baboons(,) (the gorillas) ate some bananas.

i. x-x-x story apiece juicy(,) morass expend codify(,) (hasn't enjoy) penny movie enact.

ii. x-x-x book each fresh(,) swamp spend decrypt(,) (has enliven) coin film disable.

iii. x-x-x book each fresh(,) swamp squander decrypt(,) (has enliven) coin film disable.

(31) After the President visited(,) the beleaguered senator(,) (the opponent) won the election.

i. x-x-x planet Exaggerate resent(,) kiwi deescalate henceforth(,) (hustle strengthen) enjoy wasn't include.

ii. x-x-x moon Fabricate terrorize(,) grape encourage anymore(,) (sprint endanger) dig was consider.

iii. x-x-x moon Fabricate terrorize(,) grape encourage subsequent(,) (sprint endanger) dig was consider.

(32) After the mugger attacked(,) the naive tourists(,) (the family) feared the city.

i. x-x-x device believe mental(,) warthog reimburse yours(,) (nostril destroy) giraffe hasn't when.

ii. x-x-x tool strengthen adrift(,) hog repay yourself(,) (spine deaden) sloth has ago.

iii. x-x-x tool strengthen adrift(,) hog repay someone's(,) (spine deaden) sloth has later.

(33) While the detectives investigated(,) the suspicious banker(,) (the accountant) destroyed the records on the computer.

i. x-x-x circle specifies ludicrous(,) thicket organize their(,) (elbow translate) utilize ocean would spirit portray besmirch.

ii. x-x-x square considers unreasonable(,) bush restructure themselves(,) (jab devalue) employ sea wouldn't soul say befuddle.

iii. x-x-x square considers unreasonable(,) bush restructure every(,) (jab devalue) employ sea wouldn't soul say befuddle.

(34) After the artist painted(,) the beautiful model(,) (the pair) smoked a cigarette in the garden.

i. x-x-x deny endow optimal(,) ratio however agrees(,) (unsee moisten) escape agent confide myself lemon securely.

ii. x-x-x melt donate foremost(,) rate anyway freshens(,) (see baste) flee spy coexist me lime surely.

iii. x-x-x melt donate foremost(,) rate anyway freshens(,) (see baste) flee spy coexist me lime surely.

(35) While the assistant observed(,) the young actor(,) (the understudy) rehearsed the lines with his friends.

i. x-x-x jacket recuperate reception(,) curtain enjoy hates(,) (elbow peruses) gadget sentry yourself yellow bargain either.

ii. x-x-x coat recycle receipt(,) screen have likens(,) (ear energizes) device guard yours green sale both.

iii. x-x-x coat recycle receipt(,) screen have likens(,) (ear energizes) device guard yours green sale both.

(36) While the director filmed(,) the sleepy actress(,) (the extras) drank some coffee behind the camera.

i. x-x-x lizard anyhow ritual(,) contest hence forgives(,) (tummy concur) giraffe devour exists happen unit just.

ii. x-x-x snake moreover rite(,) match therefore whitens(,) (brain conquer) sloth eat suffer occur link only.

iii. x-x-x snake moreover rite(,) match therefore whitens(,) (brain conquer) sloth eat suffer occur link only.

(37) As the lawyer contemplated(,) the shifty defendant(,) (the judge) watched the trial with growing impatience.

i. x-x-x detect destroy execution(,) siesta confess wakens(,) (talon yourself) lizard deprive wouldn't tiger depend enlarged.

ii. x-x-x feel whiten symbolism(,) nap hearten awakens(,) (beak yours) snake strip would cub glisten inflated.

iii. x-x-x feel whiten symbolism(,) nap hearten awakens(,) (beak yours) snake strip would cub glisten inflated.

(38) Because Tim teased(,) his younger sister(,) (their mother) admonished her son for his behavior.

i. x-x-x Obtain utmost(,) gather alleviate destruction(,) (basis humidify) waken decade ergo torture atlas noticed.

ii. x-x-x Get most(,) meet soften damage(,) (base dampen) awaken year thus pain map examined.

iii. x-x-x Get most(,) meet soften damage(,) (base dampen) awaken year thus pain map examined.

(39) While Sam counted(,) the small children(,) (the chaperones) boarded the bus outside the school.

i. x-x-x Gobble demeanor(,) cabin into cleans(,) (rodent dishearten) examines repay yourself equity carry only.

ii. x-x-x Gulp manner(,) house with reddens(,) (mouse aggravate) lightens pay you justice go just.

iii. x-x-x Gulp manner(,) house with reddens(,) (mouse aggravate) lightens pay you avail go just.

(40) After the athlete wrestled(,) the angry opponent(,) (the teammate) threw a punch at the referee.

i. x-x-x didn't pertain impugn(,) unsaw could brightens(,) (finger awakens) message shouldn't below lady taxi eighteen.

ii. x-x-x don't liven brighten(,) saw couldn't disheartens(,) (lip wakens) text should for lord cab seventeen.

iii. x-x-x don't liven brighten(,) saw couldn't disheartens(,) (lip wakens) text should for lord cab twenty-one.

(41) While the swordsman fought(,) the adept warrior(,) (the rogue) attacked the

archer with a dagger.

i. x-x-x didn't confesses lonely(,) emit into gives(,) (applaud before) astonish haven't verify purple into there.

ii. x-x-x don't moistens grim(,) vent among donates(,) (clap from) believe have hearten red to hereby.

iii. x-x-x don't moistens grim(,) vent among donates(,) (clap from) believe have hearten red to hereby.

(42) As the coach observed(,) the skillful player(,) (the quarterback) threw the ball to his teammate.

i. x-x-x spirit yourself sunshine(,) warthog abdicate forgets(,) (kitten represent) haven't hasn't against bottle ago around.

ii. x-x-x ghoul yours bamboo(,) hog hasten ripens(,) (cat constitute) have has from jug since during.

iii. x-x-x ghoul yours bamboo(,) hog hasten ripens(,) (cat constitute) have has from jug since during.

(43) While Anna dressed(,) the cute baby(,) (the dog) took a nap on the bed.

i. x-x-x Halt finish(,) flower ahead listens(,) (shatter after) haven't unsay until desire without neither.

ii. x-x-x Finish cease(,) plant for hears(,) (break from) have say by hope with nor.

iii. x-x-x Finish cease(,) plant for listens(,) (break from) have say by hope with nor.

(44) While the puppy sniffed(,) the fluffy kitten(,) (the bunny) pushed the ball with its nose.

i. x-x-x story from divide(,) pitcher revise stone(,) (ogle equip) haven't ingest whoever pretty shouldn't either.

ii. x-x-x book during halve(,) jar lighten concrete(,) (gaze furnish) have munch whom cute should both.

iii. x-x-x book during halve(,) jar lighten rubber(,) (gaze furnish) have munch whom cute should both.

(45) While the journalist followed(,) the corrupt politician(,) (the campaign) accepted some payments from the millionaire.

i. x-x-x peruse regardless anniversary(,) carrot fabricate evaluates(,) (remove accustom) regulate shouldn't whoever chilly caress assimilate.

ii. x-x-x browse anyway birthday(,) pea concoct intersperses(,) (take adjust) determine should wherefore cool feel alienate.

iii. x-x-x browse anyway birthday(,) pea concoct intersperses(,) (take adjust) determine should wherefore cool feel alienate.

(46) As the champion raced(,) the exhausted challenger(,) (the coach) released a sigh from the sidelines.

i. x-x-x frolic moreover realize(,) sofa evaluates computes(,) (remove hasn't) originate perform fingers listen doesn't seventy.

ii. x-x-x prance anyhow know(,) couch assesses calculates(,) (take has) compose sing toes hear do eighty.

iii. x-x-x prance anyhow know(,) couch assesses calculates(,) (take has) compose sing toes hear do eighty.

(47) As the audience booed(,) the unfunny comedian(,) (the manager) left the club in a hurry.

i. x-x-x locate wherever provide(,) atlas depress explore(,) (scamper locate) bargain acquire whoever myself ergo searches.

ii. x-x-x find anywhere give(,) map discourage decentralize(,) (rush situate) sale buy whom me thus seeks.

iii. x-x-x find anywhere give(,) map discourage decentralize(,) (rush situate) sale buy whom me thus searches.

(48) While the police watched(,) the passionate activists(,) (the protestors) carried their signs through the streets.

i. x-x-x inspect ostracize robust(,) banana interact attests(,) (expires assigns) has unsee during puppy undo however.

ii. x-x-x look exclude strong(,) grape socialize certifies(,) (dies classifies) hasn’t see from dog do how.

iii. x-x-x look exclude strong(,) grape socialize certifies(,) (dies classifies) hasn’t see from dog do how.

Appendix D

Experiment 5 and 6 Materials

Items (1) through (32) are from Hemforth et al. (2015), with the following changes: in item (12), *movie star* was changed to *actress*; in item (13), *basketball player* was changed to *athlete*; in item (14), *brain surgeon* was changed to *surgeon*; in item (17), *the grandfather of the tenant* in the Subject condition was changed to *the grandfather of the caretaker* to be consistent with the Object condition; in item (26), *fruit merchant* was changed to *merchant*; and in item (29), *bridegroom* was changed to *groom*. Items (33) through (40) and the foils were created for Experiments 5 and 6. The position of the complex NP (Subject, condition A, vs. Object, condition B) was crossed with the length of the relative clause (Short vs. Long).

(1) a. The son of the colonel who (tragically) died (of a stroke) wrote five books on tropical disease.

x-x-x gone lack else purify sell (waterfall) loss (jack went equity) click cent reach yes minimize totally.

b. The doctor met the son of the colonel who (tragically) died (of a stroke).

x-x-x afford nor earn gone lack else purify sell (waterfall) loss (jack went equity).

(2) a. The servant of the actress who (shamelessly) lied (at every chance) left town last Saturday.

x-x-x differ do burn portray cent (indication) levy (drop under sleep) soul rate guy opine.

b. Mr. Johnson visited the servant of the actress who (shamelessly) lied (at every chance).

x-x-x reply offense lamp differ do burn portray cent (indication) levy (drop under sleep).

(3) a. The attorney of the defendant who (always unconsciously) mumbled was questioned about personal matters.

x-x-x convince holy sun itemize fun (grow dissemination) textbook juice fellowship allow moreover depends.

b. The reporters interviewed the attorney of the defendant who (always unconsciously) mumbled.

x-x-x undergo giraffe is convince holy sun itemize fun (grow dissemination) textbook.

(4) a. The chauffeur of the millionaire who complained (vociferously about the schedule) was taunted by the hecklers in the crowd.

x-x-x verify cool marsh prohibit have whatsoever (secessionist true hear anywhere) go adheres nor is revise miss add yours.

b. John despised the chauffeur of the millionaire who complained (vociferously about the schedule).

x-x-x manicure hour verify cool marsh prohibit have whatsoever (secessionist true hear anywhere).

(5) a. The gardener of the executive who (always) whistled (in the shower) was nearly blind.

x-x-x intend see vie allocate lie (blender) zebra (me eat argue) sir happen flows.

b. Maria loved the gardener of the executive who (always) whistled (in the shower).

x-x-x buddy write intend see vie allocate lie (blender) zebra (me eat argue).

(6) a. The mechanic of the officer who (frequently) disappeared (from work) was from Alaska.

x-x-x moisten hear were percent holy (incentive) immigrate (fund to) nose link prayed.

b. John liked the mechanic of the officer who (frequently) disappeared (from work).

x-x-x craze go moisten hear were percent holy (incentive) immigrate (fund to).

(7) a. The client of the realtor who (suddenly and unexpectedly) vanished was married to the teacher's daughter.

x-x-x suffer me seem tremors say (snowball yes counterparts) terrify ear choices hear sale eighteens honestly.

b. The attorney defended the client of the realtor who (suddenly and unexpectedly) vanished.

x-x-x moreover sweater we suffer me seem tremors say (snowball yes counterparts) terrify.

(8) a. The agent of the author who quit (abruptly and without warning) had red hair and a mustache.

x-x-x decide do seem afford duty soften (ratify guys percent common) go give glad goal defund horizons.

b. Everyone knows the agent of the author who quit (abruptly and without warning).

x-x-x stapler cup decide do seem afford duty soften (ratify guys percent common).

(9) a. The brother of the visitor who (suddenly and unexpectedly) left was sick and had a high fever.

x-x-x anymore hay done knowledge purple (machine wide legalization) sale jab belt web laws nor yes flown.

b. Emily angered the brother of the visitor who (suddenly and unexpectedly) left.

x-x-x optimum we anymore hay done knowledge purple (machine wide legalization) sale.

(10) a. The student of the chemist who (often unexpectedly) fainted was understandably upset.

x-x-x anymore big see absorb jack (sorry preparedness) poppy web intermediaries forum.

b. Mr. Miller ignored the student of the chemist who (often unexpectedly) fainted.

x-x-x mining cleaner cent anymore big see absorb jack (sorry preparedness) poppy.

(11) a. The assistant of the politician who (reportedly always) stutters went on a cruise to the Bahamas.

x-x-x clarify eat have perform fun (depend recent) giraffe mug cent nail venues mind dad stirred.

b. Maria knows the assistant of the politician who (reportedly always) stutters.

x-x-x weird must clarify eat have perform fun (depend recent) giraffe.

(12) a. The surgeon of the actress who (almost always) gossips was recently mentioned in the press.

x-x-x thereof have we defend fun (hour agency) sawdust truck congress languages lady halt miles.

b. Louise recognized the surgeon of the actress who (almost always) gossips.

x-x-x congress eat thereof have we defend fun (hour agency) sawdust.

(13) a. The companion of the athlete who (allegedly often) stumbles didn't say anything to reporters.

x-x-x abolish lax done accept wash (conspire kitten) cuisine soot should republic you economies.

b. Ben interviewed the companion of the athlete who (allegedly often) stumbles.

x-x-x promotional go abolish lax done accept wash (conspire kitten) cuisine.

(14) a. The uncle of the surgeon who (only very rarely) vacationed was well known as an author.

x-x-x arise blank map where trip (hole shed hammer) marshland good have stock hate sell prices.

b. Friends liked the uncle of the surgeon who (only very rarely) vacationed.

x-x-x layers door arise blank map where trip (hole shed hammer) marshland.

(15) a. The relative of the actor who (too frequently) drank hated the cameraman.

x-x-x anywhere him grow relax fall (me evaluation) oasis crunch charge revisit.

b. The cameraman hated the relative of the actor who (too frequently) drank.

x-x-x revisit crunch charge anywhere him grow relax fall (me evaluation) oasis.

(16) a. The cousin of the tourist who (quickly and hastily) ate watched the waiter.

x-x-x contend are ways analyze yes (digitize luck conduit) dog chance nor expire.

b. The waiter watched the cousin of the tourist who (quickly and hastily) ate.

x-x-x expire chance nor contend are ways analyze yes (digitize luck conduit) dog.

(17) a. The grandfather of the caretaker who (very frequently) stank greeted the tenant.

x-x-x furthermore bulb in expire hope (heck objectives) elegy explode nor evolve.

b. The tenant greeted the grandfather of the caretaker who (very frequently) stank.

x-x-x evolve explode nor furthermore bulb in expire hope (heck objectives) elegy.

(18) a. The son of the teacher who (almost continuously) coughed teased the pupil.

x-x-x lose mean cent depends pond (annul thanksgiving) droplet seabed port breathe.

b. The pupil teased the son of the teacher who (almost continuously) coughed.

x-x-x breathe seabed port lose mean cent depends pond (annul thanksgiving) droplet.

(19) a. The advisor of the mayor who (often unwittingly) blinked was welcomed by the chancellor.

x-x-x foresee bay snort try soul (him topography) regatta dog homeland dine shut accomplish.

b. The chancellor welcomed the advisor of the mayor who (often unwittingly) blinked.

x-x-x accomplish homeland shut foresee bay snort try soul (him topography) regatta.

(20) a. The coach of the player who (silently and patiently) waited met the sponsor.

x-x-x think turn limit drown mine (alarm sea corridors) coyote nose acre expands.

b. The sponsor met the coach of the player who (silently and patiently) waited.

x-x-x expands nose acre think turn limit drown mine (alarm sea corridors) coyote.

(21) a. The priest of the monk who (regularly and intensely) prayed met the abbot.

x-x-x behave wait go were ban (emphasize tiger appraisal) inning give mop invade.

b. The abbot met the priest of the monk who (regularly and intensely) prayed.

x-x-x invade give mop behave wait go were ban (emphasize tiger appraisal) inning.

(22) a. The son of the journalist who (repeatedly and noisily) sneezed paid the investor.

x-x-x thus see ago infections ways (remember sale crustal) bauxite ford yes tolerate.

b. The interviewer paid the son of the journalist who (repeatedly and noisily) sneezed.

x-x-x tolerate ford yes thus see ago infections ways (remember sale crustal) bauxite.

(23) a. The co-worker of the manager who (almost always) smokes greeted the employee.

x-x-x achieve fly think suppose holy (start method) siren bastion map moreover.

b. The employee greeted the co-worker of the manager who (almost always) smokes.

x-x-x moreover bastion map achieve fly think suppose holy (start method) siren.

(24) a. The friend of the artist who (still occasionally) juggles yelled at Robert.

x-x-x happen tend met unlike ways (defend technologies) leave cleric fund server.

b. Robert yelled at the friend of the artist who (still occasionally) juggles.

x-x-x server fund eat happen tend met unlike ways (defend technologies) leave.

(25) a. The colleague of the professor who (foolishly and impetuously) resigned met the student.

x-x-x activate roar wish someone nor (traitor cent psychology) fortress fun cyst analyze.

b. The student met the colleague of the professor who (foolishly and impetuously) resigned.

x-x-x analyze fun cyst activate roar wish someone nor (traitor cent psychology) fortress.

(26) a. The uncle of the merchant who (permanently and persistently) chatters visited the boss.

x-x-x besmirch are fool deny fun (statistic yes vaccinations) metric barrier is seem.

b. The boss visited the uncle of the merchant who (permanently and persistently) chatters.

x-x-x seem barrier is besmirch are fool deny fun (statistic yes vaccinations) metric.

(27) a. The companion of the tramp who (suddenly and unexpectedly) fled warned the busker.

x-x-x announce us been exert heat (lifetime dad chemotherapy) dome refund are betray.

b. The busker warned the companion of the tramp who (suddenly and unexpectedly) fled.

x-x-x betray refund are announce us been exert heat (lifetime dad chemotherapy) dome.

(28) a. The friend of the pharmacist who (slowly and hesitantly) answered insulted the host.

x-x-x expect hang walk believe year (expand mom isomorph) tribune minimize say undo.

b. The host insulted the friend of the pharmacist who (slowly and hesitantly) answered.

x-x-x undo minimize say expect hang walk believe year (expand mom isomorphic) tribune.

(29) a. The brother of the groom who (often unknowingly) snores impressed the guest.

x-x-x anymore why hear renew rain (overalls vertebrates) drought renew plant arise.

b. The guest impressed the brother of the groom who (often unknowingly) snores.

x-x-x arise renew plant anymore why hear renew rain (overalls vertebrates) drought.

(30) a. The employee of the gardener who (carefully and dutifully) mowed liked the supplier.

x-x-x conclude go lung portray neck (elevate fact refresher) datum frank hear misinform.

b. The supplier liked the employee of the gardener who (carefully and dutifully) mowed.

x-x-x misinform frank hear conclude go lung portray neck (elevate fact refresher) datum.

(31) a. The father of the boy who (often soundly) slept bothered the grandfather.

x-x-x because game know nor got (enjoy scorer) basin hectares grape constitute.

b. The grandfather bothered the father of the boy who (often soundly) slept.

x-x-x constitute hectares grape because game know nor got (enjoy scorer) basin.

(32) a. The patron of the artist who (inconsiderately and insultingly) swore missed the client.

x-x-x illuminate why she ensure lock (villain sir renumbering) cloud soccer con defund.

b. The client missed the patron of the artist who (inconsiderately and insultingly) swore.

x-x-x defund soccer con illuminate why she ensure lock (villain sir renumbering) cloud.

(33) a. The son of the fisherman who (successfully) swam (across the lake) hated the neighbor.

x-x-x upon drop ago surpass loss (acknowledge) craze (accept get info) watts vote tolerate.

b. The neighbor hated the son of the fisherman who (successfully) swam (across the lake).

x-x-x tolerate watts vote upon drop ago surpass loss (acknowledge) craze (accept get info).

(34) a. The boss of the waitress who (often) drank (after work) greeted the cook.

x-x-x aspire nor she believe brush (mile) rider (march lion) wording walk rely.

b. The cook greeted the boss of the waitress who (often) drank (after work).

x-x-x rely wording walk aspire nor she believe brush (mile) rider (march lion).

(35) a. The partner of the writer who (sometimes) laughed (at inopportune moments) entertained the host.

x-x-x depend defy ask ignore him (advantage) vaccine (goal determine expect) collapse sky hear.

b. The host entertained the partner of the writer who (sometimes) laughed (at inopportune moments).

x-x-x hear collapse sky depend defy ask ignore him (advantage) vaccine (goal determine expect).

(36) a. The editor of the author who (often) struggled (with the computer) embarrassed the boss.

x-x-x afford hang else been leave (him) sanctuary (stay deny tomorrow) cholesterol marry pray.

b. The boss embarrassed the editor of the author who (often) struggled (with the computer).

x-x-x pray cholesterol marry afford hang else been leave (him) sanctuary (stay deny tomorrow).

(37) a. The girlfriend of the actor who (purposely) fasted (before the party) approached the publicist.

x-x-x everywhere move pin try want (mythologize) insult (review front yours) eulogize go absorb.

b. The publicist approached the girlfriend of the actor who (purposely) fasted (before the party).

x-x-x absorb eulogize go everywhere move pin try want (mythologize) insult (review front yours).

(38) a. The colleague of the librarian who (rudely) yawned (during the meeting) saw the student.

x-x-x salivate are why bandwidth train (airport) wake (agency guys anymore) feel gem anybody.

b. The student saw the colleague of the librarian who (rudely) yawned (during the meeting).

x-x-x anybody feel gem salivate are why bandwidth train (airport) wake (agency guys anymore).

(39) a. The guide of the hiker who (accidentally) fell (down the hill) called the ranger.

x-x-x compete nor hear ignore yes (disable) sale (fund go info) forget see modify.

b. The ranger called the guide of the hiker who (accidentally) fell (down the hill).

x-x-x modify forget see compete nor hear ignore yes (disable) sale (fund go info).

(40) a. The intern of the politician who (abruptly) departed (from the meeting) avoided the reporter.

x-x-x inform leaf wish accelerate luck (affirm) capture (done per percent) flavor yes moreover.

b. The reporter avoided the intern of the politician who (abruptly) departed (from the meeting).

x-x-x moreover flavor yes inform leaf wish accelerate luck (affirm) capture (done per percent).

Appendix E

Experiment 7 and 10 Materials

The attachment site required by the reflexive (Ambiguous, condition A, vs. Low, condition B, vs. High, condition C) was crossed with the length of NP2 (Short vs. Long). Items (25), (28), (29), (30), and (41) were excluded from analysis due to typos in at least one of the conditions; these typos are preserved here, such that the following materials reflect the items as presented to participants.

(1) a. The son of the (especially murderous) hunter who educated himself over the summer was successful in the first hunt.

b. The sons of the (especially murderous) hunter who educated himself over the summer were successful in the first hunt.

c. The son of the (especially murderous) hunters who educated himself over the summer was successful in the first hunt.

x-x-x gone bot done (prophesize denigrate) touches six freshman conclusion trial okay unless odd federation wall than thank were.

(2) a. The kid of the (somewhat negligent) officer who compared himself to the others was watching the news.

b. The kids of the (somewhat negligent) officer who compared himself to the others were watching the news.

c. The kid of the (somewhat negligent) officers who compared himself to the others was watching the news.

x-x-x gone two vent (believe misinform) behave holy republic invitation ad up under loss myself glad else.

(3) a. The aunt of the (awfully critical) nun who examined herself in the mirror was dissatisfied with her appearance.

b. The aunts of the (awfully critical) nun who examined herself in the mirror were dissatisfied with their appearance.

c. The aunt of the (awfully critical) nuns who examined herself in the mirror was dissatisfied with her appearance.

x-x-x swam jaw ago (perceive transmit) lap ilk behavior lieutenant air guys gotten bee governorship belt under represents.

(4) a. The aide of the (relatively forgettable) actress who enjoyed herself at

the premiere had forgotten to lock the car.

b. The aides of the (relatively forgettable) actress who enjoyed herself at the premiere had forgotten to lock the car.

c. The aide of the (relatively forgettable) actresses who enjoyed herself at the premiere had forgotten to lock the car.

x-x-x aloud luck mat (besmirch frighten) bounds fun mayor dictionary it kid freedom non southwest cent wars miss whom.

(5) a. The aide of the (very passionate) teacher who belittled herself to the supervisors was terribly humiliated.

b. The aides of the (very passionate) teacher who belittled herself to the supervisors were terribly humiliated.

c. The aide of the (very passionate) teachers who belittled herself to the supervisors was terribly humiliated.

x-x-x those mint lend (emphasize feel) involves elk luminance kilometers work mint evaporation green mushroom conjecture.

(6) a. The son of the (incredibly serious) nomad who soothed himself after the fight had proven himself to the tribe.

b. The sons of the (incredibly serious) nomad who soothed himself after the fight had proven themselves to the tribe.

c. The son of the (incredibly serious) nomads who soothed himself after the fight had proven himself to the tribe.

x-x-x nor pay else (testify amaze) sores okay operant suggestion click guy aware ads extent article owl gets bacon.

(7) a. The priest of the (truly benevolent) god who prided himself on his abilities was distracted during the ceremony.

b. The priests of the (truly benevolent) god who prided himself on his abilities were distracted during the ceremony.

c. The priest of the (truly benevolent) gods who prided himself on his abilities was distracted during the ceremony.

x-x-x ignore fun hurt (retain signify) onto dam lingua assessment top aware favoring guys newsletter length orb belonged.

(8) a. The friend of the (really sorrowful) husband who consoled himself after the incident was angry about the betrayal.

b. The friends of the (really sorrowful) husband who consoled himself after the incident were angry about the betrayal.

c. The friend of the (really sorrowful) husbands who consoled himself after the incident was angry about the betrayal.

x-x-x anyway me done (know detain) accepts hell heraldic federation click spy situated up arena beach map however.

(9) a. The friend of the (unusually difficult) actor who criticized himself at every chance quit his job in a hurry.

b. The friends of the (unusually difficult) actor who criticized himself at every chance quit their jobs in a hurry.

c. The friend of the (unusually difficult) actors who criticized himself at every chance quit his job in a hurry.

x-x-x anyway hen nor (whether however) forth lamb correction conclusion dad knows amuse bulk go none tree if nails.

(10) a. The chef of the (particularly arrogant) actor who disadvantaged himself in the competition had lost the second round.

b. The chefs of the (particularly arrogant) actor who disadvantaged himself in the competition had lost the second round.

c. The chef of the (particularly arrogant) actors who disadvantaged himself in the competition had lost the second round.

x-x-x borne luck mist (sustain neither) unlike holy defenestration conclusion arm glad distinguish sad have jump myself refer.

(11) a. The mate of the (extremely loveable) sailor who drowned himself in the ocean had bad memories from his past.

b. The mates of the (extremely loveable) sailor who drowned himself in the ocean had bad memories from their past.

c. The mate of the (extremely loveable) sailors who drowned himself in the ocean had bad memories from his past.

x-x-x swim hell shut (assail whichever) footing true movie television soul fit toward fun pin skating holy award cent.

(12) a. The guest of the (awfully humorous) prince who let himself into the hall hoped to stay for a long while.

b. The guests of the (awfully humorous) prince who let himself into the hall hoped to stay for a long while.

c. The guest of the (awfully humorous) princes who let himself into the hall hoped to stay for a long while.

x-x-x drank why sure (choose furthermore) thereof goal yam definition shy rug sit pulse diet well bill web even custom.

(13) a. The pals of the (especially confident) students who embarrassed themselves at the dance were ready to go home.

b. The pal of the (especially confident) students who embarrassed themselves at the dance was ready to go home.

c. The pals of the (especially confident) student who embarrassed themselves at the dance were ready to go home.

x-x-x fell ton done (conceive suffice) whatever glad battlefield percentage case fin smiled hat exist lose nor fund.

(14) a. The guides of the (exceptionally emotional) hikers who crippled themselves after a fall were traumatized for weeks afterwards.

b. The guide of the (exceptionally emotional) hikers who crippled themselves after a fall was traumatized for weeks afterwards.

c. The guides of the (exceptionally emotional) hiker who crippled themselves after a fall were traumatized for weeks afterwards.

x-x-x behave tall nor (deliver therefore) swam fun downfall referendum photo hear upon issue discography act basis regulation.

(15) a. The aunts of the (incredibly reverent) bishops who blamed themselves for the accident were worried about the consequences.

b. The aunt of the (incredibly reverent) bishops who blamed themselves for the accident was worried about the consequences.

c. The aunts of the (incredibly reverent) bishop who blamed themselves for the accident were worried about the consequences.

x-x-x any eve ago (contain whether) drank pain where impression the map includes else formula coast elf participated.

(16) a. The kids of the (moderately likeable) parents who introduced them-

selves at the party were too shy to socialize.

b. The kid of the (moderately likeable) parents who introduced themselves at the party was too shy to socialize.

c. The kids of the (moderately likeable) parent who introduced themselves at the party were too shy to socialize.

x-x-x gone to done (furthermore publish) toward bag resolution surrounded eld druid aware true cup beds cell gestation.

(17) a. The foes of the (slightly fraudulent) lawyers who locked themselves in the office had lost the opportunity for a promotion.

b. The foe of the (slightly fraudulent) lawyers who locked themselves in the office had lost the opportunity for a promotion.

c. The foes of the (slightly fraudulent) lawyer who locked themselves in the office had lost the opportunity for a promotion.

x-x-x say nor fund (manage wherever) destine fun sunset kilometers map down anyone sad can gift importantly moth not variables.

(18) a. The stars of the (quite remarkable) directors who outdid themselves during the play were praised during the premiere.

b. The star of the (quite remarkable) directors who outdid themselves during the play was praised during the premiere.

c. The stars of the (quite remarkable) director who outdid themselves during the play were praised during the premiere.

x-x-x yours cap fill (why scandalize) requires ill marsh federation county try even dew capitol search lest withdrew.

(19) a. The guards of the (mostly sensible) inmates who hated themselves after the catastrophe were placed under supervision for a month.

b. The guard of the (mostly sensible) inmates who hated themselves after the catastrophe was placed under supervision for a month.

c. The guards of the (mostly sensible) inmate who hated themselves after the catastrophe were placed under supervision for a month.

x-x-x suffer me sure (conjure henceforth) evolve sky dolls overwhelm album to facilitates kid format tomb understands web ago gives.

(20) a. The clerks of the (highly ethical) judges who found themselves in financial trouble had completed the paperwork incorrectly.

b. The clerk of the (highly ethical) judges who found themselves in financial trouble had completed the paperwork incorrectly.

c. The clerks of the (highly ethical) judge who found themselves in financial trouble had completed the paperwork incorrectly.

x-x-x edged if bet (peruse configure) toward fun lake generation law consider totally coy notebook oak descended pedestrians.

(21) a. The scribes of the (comparatively mystical) prophets who amused themselves during each session lost track of their work while writing.

b. The scribe of the (comparatively mystical) prophets who amused themselves during each session lost track of their work while writing.

c. The scribes of the (comparatively mystical) prophet who amused themselves during each session lost track of their work while writing.

x-x-x ours pit mom (recommend whichever) confuse top walnut foundation cancer while devices luck filed glad feels gets loose percent.

(22) a. The beaus of the (very delicate) dancers who trusted themselves during the show had planned a surprise party.

b. The beau of the (very delicate) dancers who trusted themselves during the show had planned a surprise party.

c. The beaus of the (very delicate) dancer who trusted themselves during the show had planned a surprise party.

x-x-x been glad we (cohere whomever) thereof luck pillow referendum minute pick thus wilt emperor guy contains items.

(23) a. The aides of the (relatively merciless) generals who barricaded themselves in the building had retreated to a safer location.

b. The aide of the (relatively merciless) generals who barricaded themselves in the building had retreated to a safer location.

c. The aides of the (relatively merciless) general who barricaded themselves in the building had retreated to a safer location.

x-x-x aloud when fun (desecrate shouldn't) moreover ways racecourse suggestion time loop yourself me benchmark soul kid alike altogether.

(24) a. The lords of the (quite conservative) peasants who mortified themselves at every banquet addressed the guests.

b. The lord of the (quite conservative) peasants who mortified themselves at every banquet addressed the guests.

c. The lords of the (quite conservative) peasant who mortified themselves at every banquet addressed the guests.

x-x-x arise top seen (could validate) forwards dad syndrome resolution your while manuals resistant miss expand.

(25) a. The pets of the (rather considerate) schoolgirls who behaved themselves at the park was rewarded with treats afterwards.

b. The pet of the (rather considerate) schoolgirls who behaved themselves at the park was rewarded with treats afterwards.

c. The pets of the (rather considerate) schoolgirl who behaved themselves at the park was rewarded with treats afterwards.

x-x-x told nor shut (portray moreover) emphasizing true placebo foundation hour lady whom glad timeline maze before accordance.

(26) a. The peers of the (distinctly articulate) classmates who praised themselves for the work prepared for a class presentation.

b. The peer of the (distinctly articulate) classmates who praised themselves for the work prepared for a class presentation.

c. The peers of the (distinctly articulate) classmate who praised themselves for the work prepared for a class presentation.

x-x-x try tans clue (wherever exaggerate) separates nice essence subsidiary sale alone else freedom miss guy yours temperatures.

(27) a. The guides of the (truly compassionate) prophets who lost themselves in their thoughts prayed every day before sleeping.

b. The guide of the (truly compassionate) prophets who lost themselves in their thoughts prayed every day before sleeping.

c. The guides of the (truly compassionate) prophet who lost themselves in their thoughts prayed every day before sleeping.

x-x-x pray lap runs (whether investigate) evaluate gas bulb furthermore car click bloomed litres doubt shut review overwhelm.

(28) a. The maids of the (slightly approachable) executives who calmed them-

selves after the tragedy was anxious on their first day.

b. The maid of the (slightly approachable) executives who calmed themselves after the tragedy was anxious on their first day.

c. The maids of the (slightly approachable) executive who calmed themselves after the tragedy was anxious on their first day.

x-x-x displace few they (would tranquilize) terrify fun whence correction super map defines dim tissues pick knows yours else.

(29) a. The vets of the (awfully lovable) animals who injured themselves during the procedure was rushing through appointments all day.

b. The vet of the (awfully lovable) animals who injured themselves during the procedure was rushing through appointments all day.

c. The vets of the (awfully lovable) animal who injured themselves during the procedure was rushing through appointments all day.

x-x-x unto glad drop (couldn't regale) imagine false deficit appreciate advice turn democrats bull elbow economy entrepreneur cold fact.

(30) a. The guests of the (comparatively militant) leaders who asserted themselves at the meeting was not pleased with current policies.

b. The guest of the (comparatively militant) leaders who asserted themselves at the meeting was not pleased with current policies.

c. The guests of the (comparatively militant) leader who asserted themselves at the meeting was not pleased with current policies.

x-x-x choose odds gut (wherever shouldn't) anymore your facade restaurant trip why follows nice map crystal bill century annoying.

(31) a. The king of the (especially serious) swordsman who protected himself after the attack hired more guards right away.

b. The queen of the (especially serious) swordsman who protected himself after the attack hired more guards right away.

c. The queen of the (especially serious) swordsman who protected herself after the attack hired more guards right away.

x-x-x swam true fit (whoever protect) anymore dad developer whatsoever glad win anyway sunny into affair month mean.

(32) a. The guest of the (remarkably amateur) actor who treated himself at the festival preferred listening to audio books.

b. The mom of the (remarkably amateur) actor who treated himself at the festival preferred listening to audio books.

c. The mom of the (remarkably amateur) actor who treated herself at the festival preferred listening to audio books.

x-x-x surveil luck off (anymore marry) toward fun billion assessment net done whenever synthetic catered hurt ought yours.

(33) a. The wife of the (highly capable) seamstress who entertained herself at the gallery was willing to pay a high price.

b. The son of the (highly capable) seamstress who entertained herself at the gallery was willing to pay a high price.

c. The son of the (highly capable) seamstress who entertained himself at the gallery was willing to pay a high price.

x-x-x afford low stink (whether persist) critiqued goal saxophone blackboard far guys cocaine bad carrier done hot tilt came learn.

(34) a. The pal of the (remarkably terrible) waiter who forgave himself for a mistake had just moved out of town.

b. The niece of the (remarkably terrible) waiter who forgave himself for a mistake had just moved out of town.

c. The niece of the (remarkably terrible) waiter who forgave herself for a mistake had just moved out of town.

x-x-x anymore out day (establishes stupefy) versa loss stretch conference port map updates old end saint term ear miss.

(35) a. The bride of the (particularly excellent) actress who commended herself at the performance was interviewed by a prominent journalist.

b. The dad of the (particularly excellent) actress who commended herself at the performance was interviewed by a prominent journalist.

c. The dad of the (particularly excellent) actress who commended himself at the performance was interviewed by a prominent journalist.

x-x-x relies ant seem (disintegrate anyways) conclude ash importer indication port kid forever fat degradation lord wish abilities dimensions.

(36) a. The niece of the (really elegant) duchess who welcomed herself to the wedding volunteered at the local food bank.

b. The groom of the (really elegant) duchess who welcomed herself to the wedding volunteered at the local food bank.

c. The groom of the (really elegant) duchess who welcomed himself to the wedding volunteered at the local food bank.

x-x-x remind ego done (theirs clarify) heard egg exporter subsidiary hen ivy ciphers duplication miss guy aware hear hate.

(37) a. The friend of the (highly competitive) schoolboy who challenged himself at the gym was training for a race.

b. The aunt of the (highly competitive) schoolboy who challenged himself at the gym was training for a race.

c. The aunt of the (highly competitive) schoolboy who challenged herself at the gym was training for a race.

x-x-x whatever nice had (couldn't satisfies) browse gum complexity percentage hut done ton soul consists miss ago tons.

(38) a. The boss of the (truly militant) commander who framed himself during the scandal was known for being arrogant.

b. The bride of the (truly militant) commander who framed himself during the scandal was known for being arrogant.

c. The bride of the (truly militant) commander who framed herself during the scandal was known for being arrogant.

x-x-x afford glad done (perplex furthermore) haven't holy capita undertaken moment hear dignity him exist port shall arteries.

(39) a. The wife of the (usually professional) saleswoman who disparaged herself during a meeting was rushing to make a deadline.

b. The dad of the (usually professional) saleswoman who disparaged herself during a meeting was rushing to make a deadline.

c. The dad of the (usually professional) saleswoman who disparaged himself during a meeting was rushing to make a deadline.

x-x-x speak holy eyes (petrify recollect) plenty duty reflector conclusion agency ion amazing bog custard nag must mop performs.

(40) a. The host of the (remarkably responsible) prince who rewarded himself for the success worked every weekend.

b. The maid of the (remarkably responsible) prince who rewarded himself for the success worked every weekend.

c. The maid of the (remarkably responsible) prince who rewarded herself for the success worked every weekend.

x-x-x afford out pig (deconstruct however) pushing pie carburetor silhouette lot than anymore square guess suppose.

(41) a. The pal of the (particularly talkative) schoolboy who convinced himself during the sermon was concerned with church attendance.

b. The mom of the (particularly talkative) schoolboy who convinced himself during the sermon was concerned with church attendance.

c. The mom of the (particularly talkative) schoolboy who convinced himself during the sermon was concerned with church attendance.

x-x-x contain oat may (denigrate replenishes) teleport fun prospect percentage effect pick popped lip boyfriend base happen remarkable.

(42) a. The aunt of the (rather fabulous) princess who dressed herself for the ceremony made a new decree.

b. The son of the (rather fabulous) princess who dressed herself for the ceremony made a new decree.

c. The son of the (rather fabulous) princess who dressed himself for the ceremony made a new decree.

x-x-x perspire lot jut (shouldn't classify) powering lid tablet investment jug hip circumvent hind irk get lastly.

(43) a. The friend of the (somewhat dependable) fisherman who bathed himself in the lake went fishing every month.

b. The niece of the (somewhat dependable) fisherman who bathed himself in the lake went fishing every month.

c. The niece of the (somewhat dependable) fisherman who bathed herself in the lake went fishing every month.

x-x-x relied gig ways (adore whichever) astern ink filler conclusion joy cut trust trees affects built aware.

(44) a. The aunt of the (truly beautiful) seamstress who hid herself at the convention was afraid of speaking to strangers.

b. The dad of the (truly beautiful) seamstress who hid herself at the convention was afraid of speaking to strangers.

c. The dad of the (truly beautiful) seamstress who hid himself at the convention was afraid of speaking to strangers.

x-x-x have luck else (hasn't recommend) horizons its char assessment fox wag responding yolk muscle vow consists sun battalion.

(45) a. The mate of the (fairly muscular) serviceman who celebrated himself at the reception liked to tell long stories.

b. The mom of the (fairly muscular) serviceman who celebrated himself at the reception liked to tell long stories.

c. The mom of the (fairly muscular) serviceman who celebrated herself at the reception liked to tell long stories.

x-x-x ate ten walk (abhor anymore) download rib guideline understood pug hum discusses frank bit bad term finance.

(46) a. The boss of the (extremely incompetent) businessman who paid himself from the settlement received a lot of media attention.

b. The maid of the (extremely incompetent) businessman who paid himself from the settlement received a lot of media attention.

c. The maid of the (extremely incompetent) businessman who paid herself from the settlement received a lot of media attention.

x-x-x goes ads own (demolish presuppose) deserved soot pope surrounded site won everywhere republic off our eyes mouth currently.

(47) a. The host of the (relatively liberal) councilman who impressed himself

at the reception spent many hours preparing.

b. The maid of the (relatively liberal) councilman who impressed himself at the reception spent many hours preparing.

c. The maid of the (relatively liberal) councilman who impressed herself at the reception spent many hours preparing.

x-x-x arise old ski (furthermore decompose thereof luck election federation letter pan somewhere weird lord scale youngest.

(48) a. The niece of the (incredibly diligent) waitress who chided herself over the blunder had just turned seventeen.

b. The son of the (incredibly diligent) waitress who chided herself over the blunder had just turned seventeen.

c. The son of the (incredibly diligent) waitress who chided himself over the blunder had just turned seventeen.

x-x-x admire soy glue (reconnect suppose) shading art apology contractor belt bug clocked hear big growth exceeding.

(49) a. The mate of the (really ignorant) policeman who hurt himself on the bicycle went to see the doctor.

b. The aunt of the (really ignorant) policeman who hurt himself on the bicycle went to see the doctor.

c. The aunt of the (really ignorant) policeman who hurt herself on the bicycle went to see the doctor.

x-x-x rely nor done (hence revenues) decompose art tech billion hour lady enhance glad date we mind allows.

(50) a. The mom of the (exceptionally talented) ballerina who complimented herself in an interview was bothered by the reporter.

b. The son of the (exceptionally talented) ballerina who complimented herself in an interview was bothered by the reporter.

c. The son of the (exceptionally talented) ballerina who complimented himself in an interview was bothered by the reporter.

x-x-x upon me lake (anymore displace) believe ski forecaster goggle sale nor continues them plastics beat gap decimate.

(51) a. The maid of the (rather provocative) hostess who burned herself in the kitchen was usually very careful.

b. The dad of the (rather provocative) hostess who burned herself in the kitchen was usually very careful.

c. The dad of the (rather provocative) hostess who burned himself in the kitchen was usually very careful.

x-x-x ate them been (interrupt fertilizers) prolong say twirl billion map god suppose hate percent soot revenue.

(52) a. The boss of the (fairly conservative) general who sacrificed himself for the cause was the subject of the biography.

b. The wife of the (fairly conservative) general who sacrificed himself for the cause was the subject of the biography.

c. The wife of the (fairly conservative) general who sacrificed herself for the cause was the subject of the biography.

x-x-x gone us hay (imply mountainside) anyway nor milestone imagine day zap heard been have discuss hear than integrate.

(53) a. The boss of the (quite competitive) salesman who bankrupted himself with the deal still made risky investments.

b. The wife of the (quite competitive) salesman who bankrupted himself with the deal still made risky investments.

c. The wife of the (quite competitive) salesman who bankrupted herself with the deal still made risky investments.

x-x-x were nog can (hence consumption) devote vouch magnitude improve wind mop see fresh hour haven't distinctive.

(54) a. The mate of the (moderately competent) surgeon who prepared himself for the worst walked into the operating room.

b. The niece of the (moderately competent) surgeon who prepared himself for the worst walked into the operating room.

c. The niece of the (moderately competent) surgeon who prepared herself for the worst walked into the operating room.

x-x-x have me sink (anyhow reclaims) justify sail cupcake observe ant ago those dismiss lust sell ourselves sent.

(55) a. The mom of the (unusually popular) schoolgirl who helped herself to a dessert was waiting at home.

b. The dad of the (unusually popular) schoolgirl who helped herself to a dessert was waiting at home.

c. The dad of the (unusually popular) schoolgirl who helped himself to a dessert was waiting at home.

x-x-x this us deny (indicate windows) subtract clap lesson million nor has vectors maze percent may rule.

(56) a. The maid of the (distinctly glamorous) princess who bewildered herself at the party had spilled the drinks.

b. The priest of the (distinctly glamorous) princess who bewildered herself at the party had spilled the drinks.

c. The priest of the (distinctly glamorous) princess who bewildered himself at the party had spilled the drinks.

x-x-x rely fax we (predicts pacemaker) divest goal fireman library me why hence bay should tour submit.

(57) a. The boss of the (somewhat considerate) mailman who cut himself on a nail was waiting for the doctor.

b. The wife of the (somewhat considerate) mailman who cut himself on a nail was waiting for the doctor.

c. The wife of the (somewhat considerate) mailman who cut herself on a nail was waiting for the doctor.

x-x-x has nice too (wherever testimonies) surpass rice us imagine say if tens guys nations self with these.

(58) a. The friend of the (slightly cynical) businessman who invited himself to the party hated small children.

b. The niece of the (slightly cynical) businessman who invited himself to the party hated small children.

c. The niece of the (slightly cynical) businessman who invited herself to the party hated small children.

x-x-x alike them been (surveil download) preregister door eagle percent me than among pulse click achieved.

(59) a. The pal of the (mostly responsible) janitor who taught himself in the

mornings loved going to the theatre.

b. The wife of the (mostly responsible) janitor who taught himself in the mornings loved going to the theatre.

c. The wife of the (mostly responsible) janitor who taught herself in the mornings loved going to the theatre.

x-x-x this spam shut (happen conclusions) persist gap domain fridge cent do imbibe blame month wish sum haunted.

(60) a. The pal of the (very compassionate) fireman who admired himself for his charity was friends with many activists.

b. The mom of the (very compassionate) fireman who admired himself for his charity was friends with many activists.

c. The mom of the (very compassionate) fireman who admired herself for her charity was friends with many activists.

x-x-x which tin redo (upon interruptions) celebrate holy solution dumpster we than enlist have should skirt deny surface.

Appendix F

Experiment 8 Materials

The attachment site required by the reflexive (Ambiguous, condition A, vs. Low, condition B, vs. High, condition C) was crossed with the length of NP1 (Short vs. Long).

(1) a. The (especially confident) son of the hunter who educated himself over the summer was successful in the first hunt.

b. The (especially confident) sons of the hunter who educated himself over the summer were successful in the first hunt.

c. The (especially confident) son of the hunters who educated himself over the summer was successful in the first hunt.

x-x-x (prophesize denigrate) gone bot done touches six freshman conclusion trial okay unless odd federation wall than thank were.

(2) a. The (somewhat serious) kid of the officer who compared himself to the others was watching the news.

b. The (somewhat serious) kids of the officer who compared himself to the others were watching the news.

c. The (somewhat serious) kid of the officers who compared himself to the others was watching the news.

x-x-x (believe misinform) gone two vent behave holy republic invitation ad up under loss myself glad else.

(3) a. The (awfully critical) aunt of the nun who examined herself in the mirror was dissatisfied with her appearance.

b. The (awfully critical) aunts of the nun who examined herself in the mirror were dissatisfied with their appearance.

c. The (awfully critical) aunt of the nuns who examined herself in the mirror was dissatisfied with her appearance.

x-x-x (perceive transmit) swam jaw ago lap ilk behavior lieutenant air guys gotten bee governorship belt under represents.

(4) a. The (relatively dependable) aide of the actress who enjoyed herself at the premiere had forgotten to lock the car.

b. The (relatively dependable) aides of the actress who enjoyed herself at the premiere had forgotten to lock the car.

c. The (relatively dependable) aide of the actresses who enjoyed herself at the premiere had forgotten to lock the car.

x-x-x (besmirch frighten) aloud luck mat bounds fun mayor dictionary it kid freedom non southwest cent wars miss whom.

(5) a. The (very passionate) aide of the teacher who belittled herself to the supervisors was terribly humiliated.

b. The (very passionate) aides of the teacher who belittled herself to the supervisors were terribly humiliated.

c. The (very passionate) aide of the teachers who belittled herself to the supervisors was terribly humiliated.

x-x-x (lend emphasize) those mint feel involves elk luminance kilometers work mint evaporation green mushroom conjecture.

(6) a. The (incredibly serious) son of the nomad who soothed himself after the fight had proven himself to the tribe.

b. The (incredibly serious) sons of the nomad who soothed himself after the fight had proven themselves to the tribe.

c. The (incredibly serious) son of the nomads who soothed himself after the fight had proven himself to the tribe.

x-x-x (testify amaze) nor pay else sores okay operant suggestion click guy aware ads extent article owl gets bacon.

(7) a. The (truly benevolent) priest of the god who prided himself on his abilities was distracted during the ceremony.

b. The (truly benevolent) priests of the god who prided himself on his abilities were distracted during the ceremony.

c. The (truly benevolent) priest of the gods who prided himself on his abilities was distracted during the ceremony.

x-x-x (retain signify) ignore fun hurt onto dam lingua assessment top aware favoring guys newsletter length orb belonged.

(8) a. The (really sorrowful) friend of the husband who consoled himself after the incident was angry about the betrayal.

b. The (really sorrowful) friends of the husband who consoled himself after the incident were angry about the betrayal.

c. The (really sorrowful) friend of the husbands who consoled himself after the incident was angry about the betrayal.

x-x-x (know detain) anyway me done accepts hell heraldic federation click spy situated up arena beach map however.

(9) a. The (unusually difficult) friend of the actor who criticized himself at

every chance quit his job in a hurry.

b. The (unusually difficult) friends of the actor who criticized himself at every chance quit their jobs in a hurry.

c. The (unusually difficult) friend of the actors who criticized himself at every chance quit his job in a hurry.

x-x-x (whether however) anyway hen nor forth lamb correction conclusion dad knows amuse bulk go none tree if nails.

(10) a. The (particularly arrogant) chef of the actor who disadvantaged himself in the competition had lost the second round.

b. The (particularly arrogant) chefs of the actor who disadvantaged himself in the competition had lost the second round.

c. The (particularly arrogant) chef of the actors who disadvantaged himself in the competition had lost the second round.

x-x-x (neither unlike) borne luck mist sustain holy defenestration conclusion arm glad distinguish sad have jump myself refer.

(11) a. The (extremely loveable) mate of the sailor who drowned himself in the ocean had bad memories from his past.

b. The (extremely loveable) mates of the sailor who drowned himself in the ocean had bad memories from their past.

c. The (extremely loveable) mate of the sailors who drowned himself in the ocean had bad memories from his past.

x-x-x (assail whichever) swim hell shut footing true movie television soul fit toward fun pin skating holy award cent.

(12) a. The (awfully humorous) guest of the prince who let himself into the hall hoped to stay for a long while.

b. The (awfully humorous) guests of the prince who let himself into the hall hoped to stay for a long while.

c. The (awfully humorous) guest of the princes who let himself into the hall hoped to stay for a long while.

x-x-x (choose furthermore) drank why sure thereof goal yam definition shy rug sit pulse diet well bill web even custom.

(13) a. The (especially generous) pals of the students who embarrassed themselves at the dance were ready to go home.

b. The (especially generous) pal of the students who embarrassed themselves at the dance was ready to go home.

c. The (especially generous) pals of the student who embarrassed themselves at the dance were ready to go home.

x-x-x (conceive suffice) fell ton done whatever glad battlefield percentage case fin smiled hat exist lose nor fund.

(14) a. The (exceptionally emotional) guides of the hikers who crippled themselves after a fall were traumatized for weeks afterwards.

b. The (exceptionally emotional) guide of the hikers who crippled themselves after a fall was traumatized for weeks afterwards.

c. The (exceptionally emotional) guides of the hiker who crippled themselves after a fall were traumatized for weeks afterwards.

x-x-x (deliver therefore) behave tall nor swam fun downfall referendum photo hear upon issue discography act basis regulation.

(15) a. The (incredibly predictable) aunts of the bishops who blamed themselves for the accident were worried about the consequences.

b. The (incredibly predictable) aunt of the bishops who blamed themselves for the accident was worried about the consequences.

c. The (incredibly predictable) aunts of the bishop who blamed themselves for the accident were worried about the consequences.

x-x-x (contain whether) any eve ago drank pain where impression the map includes else formula coast elf participated.

(16) a. The (moderately likeable) kids of the parents who introduced themselves at the party were too shy to socialize.

b. The (moderately likeable) kid of the parents who introduced themselves at the party was too shy to socialize.

c. The (moderately likeable) kids of the parent who introduced themselves at the party were too shy to socialize.

x-x-x (furthermore publish) gone to done toward bag resolution surrounded eld druid aware true cup beds cell gestation.

(17) a. The (slightly fraudulent) foes of the lawyers who locked themselves in the office had lost the opportunity for a promotion.

b. The (slightly fraudulent) foe of the lawyers who locked themselves in the office had lost the opportunity for a promotion.

c. The (slightly fraudulent) foes of the lawyer who locked themselves in the office had lost the opportunity for a promotion.

x-x-x (manage wherever) say nor fund destine fun sunset kilometers map down anyone sad can gift importantly moth not variables.

(18) a. The (quite remarkable) stars of the directors who outdid themselves during the play were praised during the premiere.

b. The (quite remarkable) star of the directors who outdid themselves during the play was praised during the premiere.

c. The (quite remarkable) stars of the director who outdid themselves during the play were praised during the premiere.

x-x-x (why scandalize) yours cap fill requires ill marsh federation county try even dew capitol search lest withdrew.

(19) a. The (mostly sensible) guards of the inmates who hated themselves after the catastrophe were placed under supervision for a month.

b. The (mostly sensible) guard of the inmates who hated themselves after the catastrophe was placed under supervision for a month.

c. The (mostly sensible) guards of the inmate who hated themselves after the catastrophe were placed under supervision for a month.

x-x-x (conjure henceforth) suffer me sure evolve sky dolls overwhelm album to facilitates kid format tomb understands web ago gives.

(20) a. The (highly ethical) clerks of the judges who found themselves in financial trouble had completed the paperwork incorrectly.

b. The (highly ethical) clerk of the judges who found themselves in financial trouble had completed the paperwork incorrectly.

c. The (highly ethical) clerks of the judge who found themselves in financial trouble had completed the paperwork incorrectly.

x-x-x (peruse configure) edged if bet toward fun lake generation law consider totally coy notebook oak descended pedestrians.

(21) a. The (comparatively mystical) scribes of the prophets who amused

themselves during each session lost track of their work while writing.

b. The (comparatively mystical) scribe of the prophets who amused themselves during each session lost track of their work while writing.

c. The (comparatively mystical) scribes of the prophet who amused themselves during each session lost track of their work while writing.

x-x-x (recommend whichever) ours pit mom confuse top walnut foundation cancer while devices luck filed glad feels gets loose percent.

(22) a. The (very delicate) beaus of the dancers who trusted themselves during the show had planned a surprise party.

b. The (very delicate) beau of the dancers who trusted themselves during the show had planned a surprise party.

c. The (very delicate) beaus of the dancer who trusted themselves during the show had planned a surprise party.

x-x-x (cohere whomever) been glad we thereof luck pillow referendum minute pick thus wilt emperor guy contains items.

(23) a. The (relatively merciless) aides of the generals who barricaded themselves in the building had retreated to a safer location.

b. The (relatively merciless) aide of the generals who barricaded themselves in the building had retreated to a safer location.

c. The (relatively merciless) aides of the general who barricaded themselves in the building had retreated to a safer location.

x-x-x (desecrate shouldn't) aloud when fun moreover ways racecourse suggestion time loop yourself me benchmark soul kid alike altogether.

(24) a. The (quite conservative) lords of the peasants who mortified themselves at every banquet addressed the guests.

b. The (quite conservative) lord of the peasants who mortified themselves at every banquet addressed the guests.

c. The (quite conservative) lords of the peasant who mortified themselves at every banquet addressed the guests.

x-x-x (could validate) arise top seen forwards dad syndrome resolution your while manuals resistant miss expand.

(25) a. The (rather adorable) pets of the schoolgirls who behaved themselves at the park were rewarded with treats afterwards.

b. The (rather adorable) pet of the schoolgirls who behaved themselves at the park was rewarded with treats afterwards.

c. The (rather adorable) pets of the schoolgirl who behaved themselves at the park were rewarded with treats afterwards.

x-x-x (portray moreover) told nor shut emphasizing true placebo foundation hour lady whom glad timeline maze before accordance.

(26) a. The (distinctly studious) peers of the classmates who praised themselves for the work prepared for a class presentation.

b. The (distinctly studious) peer of the classmates who praised themselves for the work prepared for a class presentation.

c. The (distinctly studious) peers of the classmate who praised themselves for the work prepared for a class presentation.

x-x-x (wherever exaggerate) try tans clue separates nice essence subsidiary sale alone else freedom miss guy yours temperatures.

(27) a. The (truly compassionate) guides of the prophets who lost themselves in their thoughts prayed every day before sleeping.

b. The (truly compassionate) guide of the prophets who lost themselves in their thoughts prayed every day before sleeping.

c. The (truly compassionate) guides of the prophet who lost themselves in their thoughts prayed every day before sleeping.

x-x-x (whether investigate) pray lap runs evaluate gas bulb furthermore car click bloomed litres doubt shut review overwhelm.

(28) a. The (slightly approachable) maids of the executives who calmed themselves after the tragedy were anxious on their first day.

b. The (slightly approachable) maid of the executives who calmed themselves after the tragedy was anxious on their first day.

c. The (slightly approachable) maids of the executive who calmed themselves after the tragedy were anxious on their first day.

x-x-x (would tranquilize) displace few they terrify fun whence correction super map defines dim tissues pick knows yours else.

(29) a. The (awfully likeable) vets of the animals who injured themselves during the procedure were rushing through appointments all day.

b. The (awfully likeable) vet of the animals who injured themselves during the procedure was rushing through appointments all day.

c. The (awfully likeable) vets of the animal who injured themselves during the procedure were rushing through appointments all day.

x-x-x (couldn't regale) unto glad drop imagine false deficit appreciate advice turn democrats bull elbow economy entrepreneur cold fact.

(30) a. The (comparatively militant) guests of the leaders who asserted themselves at the meeting were not pleased with current policies.

b. The (comparatively militant) guest of the leaders who asserted themselves at the meeting was not pleased with current policies.

c. The (comparatively militant) guests of the leader who asserted themselves at the meeting were not pleased with current policies.

x-x-x (wherever shouldn't) choose odds gut anymore your facade restaurant trip why follows nice map crystal bill century annoying.

(31) a. The (especially serious) king of the swordsman who protected himself after the attack hired more guards right away.

b. The (especially serious) queen of the swordsman who protected himself after the attack hired more guards right away.

c. The (especially serious) queen of the swordsman who protected herself after the attack hired more guards right away.

x-x-x (whoever protect) swam true fit anymore dad developer whatsoever glad win anyway sunny into affair month mean.

(32) a. The (remarkably critical) guest of the actor who treated himself at the festival preferred listening to audio books.

b. The (remarkably critical) mom of the actor who treated himself at the festival preferred listening to audio books.

c. The (remarkably critical) mom of the actor who treated herself at the festival preferred listening to audio books.

x-x-x (anymore marry) surveil luck off toward fun billion assessment net done whenever synthetic catered hurt ought yours.

(33) a. The (highly capable) wife of the seamstress who entertained herself at

the gallery was willing to pay a high price.

b. The (highly capable) son of the seamstress who entertained herself at the gallery was willing to pay a high price.

c. The (highly capable) son of the seamstress who entertained himself at the gallery was willing to pay a high price.

x-x-x (whether persist) afford low stink critiqued goal saxophone blackboard far guys cocaine bad carrier done hot tilt came learn.

(34) a. The (remarkably ignorant) pal of the waiter who forgave himself for a mistake had just moved out of town.

b. The (remarkably ignorant) niece of the waiter who forgave himself for a mistake had just moved out of town.

c. The (remarkably ignorant) niece of the waiter who forgave herself for a mistake had just moved out of town.

x-x-x (establishes stupefy) anymore out day versa loss stretch conference port map updates old end saint term ear miss.

(35) a. The (particularly sociable) bride of the actress who commended herself at the performance was interviewed by a prominent journalist.

b. The (particularly sociable) dad of the actress who commended herself at the performance was interviewed by a prominent journalist.

c. The (particularly sociable) dad of the actress who commended himself at the performance was interviewed by a prominent journalist.

x-x-x (disintegrate anyways) relies ant seem conclude ash importer indication port kid forever fat degradation lord wish abilities dimensions.

(36) a. The (really elegant) niece of the duchess who welcomed herself to the wedding volunteered at the local food bank.

b. The (really elegant) groom of the duchess who welcomed herself to the wedding volunteered at the local food bank.

c. The (really elegant) groom of the duchess who welcomed himself to the wedding volunteered at the local food bank.

x-x-x (theirs clarify) remind ego done heard egg exporter subsidiary hen ivy ciphers duplication miss guy aware hear hate.

(37) a. The (highly competitive) friend of the schoolboy who challenged himself at the gym was training for a race.

b. The (highly competitive) aunt of the schoolboy who challenged himself at the gym was training for a race.

c. The (highly competitive) aunt of the schoolboy who challenged herself at the gym was training for a race.

x-x-x (couldn't satisfies) whatever nice had browse gum complexity percentage hut done ton soul consists miss ago tons.

(38) a. The (truly passionate) boss of the commander who framed himself during the scandal was known for being arrogant.

b. The (truly passionate) bride of the commander who framed himself during the scandal was known for being arrogant.

c. The (truly passionate) bride of the commander who framed herself during the scandal was known for being arrogant.

x-x-x (perplex furthermore) afford glad done haven't holy capita undertaken moment hear dignity him exist port shall arteries.

(39) a. The (extremely professional) wife of the saleswoman who disparaged herself during a meeting was rushing to make a deadline.

b. The (extremely professional) dad of the saleswoman who disparaged herself during a meeting was rushing to make a deadline.

c. The (extremely professional) dad of the saleswoman who disparaged himself during a meeting was rushing to make a deadline.

x-x-x (petrify recollect) speak holy eyes plenty duty reflector conclusion agency ion amazing bog custard nag must mop performs.

(40) a. The (remarkably responsible) host of the prince who rewarded himself

for the success worked every weekend.

b. The (remarkably responsible) maid of the prince who rewarded himself for the success worked every weekend.

c. The (remarkably responsible) maid of the prince who rewarded herself for the success worked every weekend.

x-x-x (deconstruct however) afford out pig pushing pie carburetor silhouette lot than anymore square guess suppose.

(41) a. The (particularly talkative) pal of the schoolboy who convinced himself during the sermon was concerned with church attendance.

b. The (particularly talkative) mom of the schoolboy who convinced himself during the sermon was concerned with church attendance.

c. The (particularly talkative) mom of the schoolboy who convinced herself during the sermon was concerned with church attendance.

x-x-x (denigrate replenishes) contain oat may teleport fun prospect percentage effect pick popped lip boyfriend base happen remarkable.

(42) a. The (rather fabulous) aunt of the princess who dressed herself for the ceremony made a new decree.

b. The (rather fabulous) son of the princess who dressed herself for the ceremony made a new decree.

c. The (rather fabulous) son of the princess who dressed himself for the ceremony made a new decree.

x-x-x (shouldn't classify) perspire lot jut powering lid tablet investment jug hip circumvent hind irk get lastly.

(43) a. The (somewhat dependable) friend of the fisherman who bathed himself in the lake went fishing every month.

b. The (somewhat dependable) niece of the fisherman who bathed himself in the lake went fishing every month.

c. The (somewhat dependable) niece of the fisherman who bathed herself in the lake went fishing every month.

x-x-x (adore whichever) relied gig ways astern ink filler conclusion joy cut trust trees affects built aware.

(44) a. The (truly generous) aunt of the seamstress who hid herself at the convention was afraid of speaking to strangers.

b. The (truly generous) dad of the seamstress who hid herself at the convention was afraid of speaking to strangers.

c. The (truly generous) dad of the seamstress who hid himself at the convention was afraid of speaking to strangers.

x-x-x (hasn't recommend) have luck else horizons its char assessment fox wag responding yolk muscle vow consists sun battalion.

(45) a. The (fairly sociable) mate of the serviceman who celebrated himself at the reception liked to tell long stories.

b. The (fairly sociable) mom of the serviceman who celebrated himself at the reception liked to tell long stories.

c. The (fairly sociable) mom of the serviceman who celebrated herself at the reception liked to tell long stories.

x-x-x (abhor anymore) ate ten walk download rib guideline understood pug hum discusses frank bit bad term finance.

(46) a. The (extremely incompetent) boss of the businessman who paid himself from the settlement received a lot of media attention.

b. The (extremely incompetent) maid of the businessman who paid himself from the settlement received a lot of media attention.

c. The (extremely incompetent) maid of the businessman who paid herself from the settlement received a lot of media attention.

x-x-x (demolish presuppose) goes ads own deserved soot pope surrounded site won everywhere republic off our eyes mouth currently.

(47) a. The (relatively liberal) host of the councilman who impressed himself at the reception spent many hours preparing.

b. The (relatively liberal) maid of the councilman who impressed himself at the reception spent many hours preparing.

c. The (relatively liberal) maid of the councilman who impressed herself at the reception spent many hours preparing.

x-x-x (furthermore decompose) arise old ski thereof luck election federation letter pan somewhere weird lord scale youngest.

(48) a. The (incredibly diligent) niece of the waitress who chided herself over the blunder had just turned seventeen.

b. The (incredibly diligent) son of the waitress who chided herself over the blunder had just turned seventeen.

c. The (incredibly diligent) son of the waitress who chided himself over the blunder had just turned seventeen.

x-x-x (reconnect suppose) admire soy glue shading art apology contractor belt bug clocked hear big growth exceeding.

(49) a. The (really ignorant) mate of the policeman who hurt himself on the bicycle went to see the doctor.

b. The (really ignorant) aunt of the policeman who hurt himself on the bicycle went to see the doctor.

c. The (really ignorant) aunt of the policeman who hurt herself on the bicycle went to see the doctor.

x-x-x (hence revenues) rely nor done decompose art tech billion hour lady enhance glad date we mind allows.

(50) a. The (exceptionally talented) mom of the ballerina who complimented herself in an interview was bothered by the reporter.

b. The (exceptionally talented) son of the ballerina who complimented herself in an interview was bothered by the reporter.

c. The (exceptionally talented) son of the ballerina who complimented himself in an interview was bothered by the reporter.

x-x-x (anymore displace) upon me lake believe ski forecaster goggle sale nor continues them plastics beat gap decimate.

(51) a. The (rather considerate) maid of the hostess who burned herself in the kitchen was usually very careful.

b. The (rather considerate) dad of the hostess who burned herself in the kitchen was usually very careful.

c. The (rather considerate) dad of the hostess who burned himself in the kitchen was usually very careful.

x-x-x (interrupt fertilizers) ate them been prolong say twirl billion map god suppose hate percent soot revenue.

(52) a. The (fairly conservative) boss of the general who sacrificed himself

for the cause was the subject of the biography.

b. The (fairly conservative) wife of the general who sacrificed himself for the cause was the subject of the biography.

c. The (fairly conservative) wife of the general who sacrificed herself for the cause was the subject of the biography.

x-x-x (imply mountainside) gone us hay anyway nor milestone imagine day zap heard been have discuss hear than integrate.

(53) a. The (quite competitive) boss of the salesman who bankrupted himself with the deal still made risky investments.

b. The (quite competitive) wife of the salesman who bankrupted himself with the deal still made risky investments.

c. The (quite competitive) wife of the salesman who bankrupted herself with the deal still made risky investments.

x-x-x (hence consumption) were nog can devote vouch magnitude improve wind mop see fresh hour haven't distinctive.

(54) a. The (moderately competent) mate of the surgeon who prepared himself for the worst walked into the operating room.

b. The (moderately competent) niece of the surgeon who prepared himself for the worst walked into the operating room.

c. The (moderately competent) niece of the surgeon who prepared herself for the worst walked into the operating room.

x-x-x (anyhow reclaims) have me sink justify sail cupcake observe ant ago those dismiss lust sell ourselves sent.

(55) a. The (unusually popular) mom of the schoolgirl who helped herself to a dessert was waiting at home.

b. The (unusually popular) dad of the schoolgirl who helped herself to a dessert was waiting at home.

c. The (unusually popular) dad of the schoolgirl who helped himself to a dessert was waiting at home.

x-x-x (indicate windows) this us deny subtract clap lesson million nor has vectors maze percent may rule.

(56) a. The (particularly sensitive) maid of the princess who bewildered herself at the party had spilled the drinks.

b. The (particularly sensitive) priest of the princess who bewildered herself at the party had spilled the drinks.

c. The (particularly sensitive) priest of the princess who bewildered himself at the party had spilled the drinks.

x-x-x (predicts pacemaker) rely fax we divest goal fireman library me why hence bay should tour submit.

(57) a. The (somewhat considerate) boss of the mailman who cut himself on a nail was waiting for the doctor.

b. The (somewhat considerate) wife of the mailman who cut himself on a nail was waiting for the doctor.

c. The (somewhat considerate) wife of the mailman who cut herself on a nail was waiting for the doctor.

x-x-x (wherever testimonies) has nice too surpass rice us imagine say if tens guys nations self with these.

(58) a. The (slightly cynical) friend of the businessman who invited himself to the party hated small children.

b. The (slightly cynical) niece of the businessman who invited himself to the party hated small children.

c. The (slightly cynical) niece of the businessman who invited herself to the party hated small children.

x-x-x (surveil download) alike them been preregister door eagle percent me than among pulse click achieved.

(59) a. The (mostly responsible) pal of the janitor who taught himself in the mornings loved going to the theatre.

b. The (mostly responsible) wife of the janitor who taught himself in the mornings loved going to the theatre.

c. The (mostly responsible) wife of the janitor who taught herself in the mornings loved going to the theatre.

x-x-x (happen conclusions) this spam shut persist gap domain fridge cent do imbibe blame month wish sum haunted.

(60) a. The (very compassionate) pal of the fireman who admired himself for his charity was friends with many activists.

b. The (very compassionate) mom of the fireman who admired himself for his charity was friends with many activists.

c. The (very compassionate) mom of the fireman who admired herself for her charity was friends with many activists.

x-x-x (upon interruptions) which tin redo celebrate holy solution dumpster we than enlist have should skirt deny surface.

Appendix G

Experiment 9 Materials

The experiment crossed the presence of modification (None vs. N1 vs. N2) with the length of the relative clause (Short vs. Long). Items (1) through (32) were based on the items from Hemforth et al. (2015), items (33) through (40) were based on the novel items from Experiments 5 and 6, and items (41) through (48) were written for Experiment 10.

(1) The doctor met the (relatively famous) son of the (relatively famous) colonel who (tragically) died (of a stroke).
x-x-x afford nor earn (implode skyscraper) gone lack else (implode skyscraper) purify sell (waterfall) loss (jack went equity).

(2) Mr. Johnson visited the (very beautiful) servant of the (very beautiful) actress who (shamelessly) lied (at every chance).

x-x-x reply offense lamp (sell revision) differ do burn (sell revision) portray cent (indication) levy (drop under sleep).

(3) The reporters interviewed the (extremely liberal) attorney of the (extremely liberal) defendant who (always unconsciously) mumbled.

x-x-x undergo giraffe is (deceive border) convince holy sun (deceive border) itemize fun (grow dissemination) textbook.

(4) Mrs. Smith despised the (rather ignorant) chauffeur of the (rather ignorant) millionaire who complained (vociferously about the schedule).

x-x-x four manicure hour (declare creature) verify cool marsh (declare creature) prohibit have whatsoever (secessionist true hear anywhere).

(5) Mrs. Jones loved the (exceptionally clever) gardener of the (exceptionally clever) executive who (always) whistled (in the shower).

x-x-x them buddy write (rhapsodize terrify) intend see vie (rhapsodize terrify) allocate lie (blender) zebra (me eat argue).

(6) Mr. Thompson liked the (very attractive) mechanic of the (very attractive) officer who (frequently) disappeared (from work).

x-x-x prize craze go (feign extradite) moisten hear were (feign extradite) percent holy (incentive) immigrate (fund to).

(7) The attorney defended the (unusually educated) client of the (unusually educated) realtor who (suddenly and unexpectedly) vanished.

x-x-x moreover sweater we (amaze football) suffer me seem (amaze football) tremors say (snowball yes counterparts) terrify.

(8) Most people know the (particularly outgoing) agent of the (particularly outgoing) author who quit (abruptly and without warning).

x-x-x they stapler cup (reminisce generate) decide do seem (reminisce generate) afford duty soften (ratify guys percent common).

(9) The maid angered the (remarkably confident) brother of the (remarkably confident) visitor who (suddenly and unexpectedly) left.

x-x-x goes optimum we (suppose strengthen) anymore hay done (suppose strengthen) knowledge purple (machine wide legalization) sale.

(10) Mr. Miller ignored the (fairly eager) student of the (fairly eager) chemist who (often unexpectedly) fainted.

x-x-x mining cleaner cent (didn't pieces) anymore big see (didn't pieces) absorb jack (sorry preparedness) poppy.

(11) The journalist knows the (highly loyal) assistant of the (highly loyal) politician who (reportedly always) stutters.

x-x-x emphasize weird must (goes forever) clarify eat have (goes forever) perform fun (depend recent) giraffe.

(12) Mrs. Thomas recognized the (fairly competent) surgeon of the (fairly competent) actress who (almost always) gossips.

x-x-x portray congress eat (deny elevate) thereof have we (deny elevate) defend fun (hour agency) sawdust.

(13) The intern interviewed the (somewhat impolite) companion of the (somewhat impolite) athlete who (allegedly often) stumbles.

x-x-x yourself promotional go (depress decide) abolish lax done (depress decide) accept wash (conspire kitten) cuisine.

(14) The friends liked the (extremely kind) uncle of the (extremely kind) surgeon who (only very rarely) vacationed.

x-x-x breathe layers door (admonish trick) arise blank map (admonish trick) where trip (hole shed hammer) marshland.

(15) The cameraman hated the (incredibly observant) relative of the (incredibly observant) actor who (too frequently) drank.

x-x-x revisit crunch charge (vanish garden) anywhere him grow (vanish garden) relax fall (me evaluation) oasis.

(16) The waiter watched the (really superficial) cousin of the (really superficial) tourist who (quickly and hastily) ate.

x-x-x expire chance nor (follow decade) contend are ways (follow decade) analyze yes (digitize luck conduit) dog.

(17) The tenant greeted the (rather friendly) grandfather of the (rather friendly) caretaker who (very frequently) stank.

x-x-x evolve explode nor (send button) furthermore bulb in (send button) expire hope (heck objectives) elegy.

(18) The pupil teased the (awfully energetic) son of the (awfully energetic) teacher who (almost continuously) coughed.

x-x-x breathe seabed port (grow courthouse) lose mean cent (grow courthouse) depends pond (annul thanksgiving) droplet.

(19) The chancellor welcomed the (moderately competitive) advisor of the (moderately competitive) mayor who (often unwittingly) blinked.

x-x-x accomplish homeland shut (enshrine denigrate) foresee bay snort (enshrine denigrate) try soul (him topography) regatta.

(20) The sponsor met the (remarkably short) coach of the (remarkably short) player who (silently and patiently) waited.

x-x-x expands nose acre (provide witchcraft) think turn limit (provide witchcraft) drown mine (alarm sea corridors) coyote.

(21) The abbot met the (unusually popular) priest of the (unusually popular) monk who (regularly and intensely) prayed.

x-x-x invade give mop (reconvene neglect) behave wait go (reconvene neglect) were ban (emphasize tiger appraisal) inning.

(22) The interviewer paid the (tremendously funny) son of the (tremendously funny) journalist who (repeatedly and noisily) sneezed.

x-x-x tolerate ford yes (synergize streetlight) thus see ago (synergize street-light) infections ways (remember sale crustal) bauxite.

(23) The employee greeted the (relatively tall) co-worker of the (relatively tall) manager who (almost always) smokes.

x-x-x moreover bastion map (expire cardigan) achieve fly think (expire cardigan) suppose holy (start method) siren.

(24) The manager scolded the (truly generous) friend of the (truly generous) artist who (still occasionally) juggles.

x-x-x belong serve fund (eat afterglow) happen tend met (eat afterglow) unlike ways (defend technologies) leave.

(25) The student met the (somewhat handsome) colleague of the (somewhat handsome) professor who (foolishly and impetuously) resigned.

x-x-x analyze fun cyst (contradict daylight) activate roar wish (contradict daylight) someone nor (traitor cent psychology) fortress.

(26) The boss visited the (slightly ugly) uncle of the (slightly ugly) merchant who (permanently and persistently) chatters.

x-x-x seem barrier is (avail web) besmirch are fool (avail web) deny fun (statistic yes vaccinations) metric.

(27) The busker warned the (moderately likeable) companion of the (moderately likeable) tramp who (suddenly and unexpectedly) fled.

x-x-x betray refund are (penalize cloud) announce us been (penalize cloud) exert heat (lifetime dad chemotherapy) dome.

(28) The host insulted the (quite skinny) friend of the (quite skinny) pharmacist who (slowly and hesitantly) answered.

x-x-x undo minimize say (those piano) expect hang walk (those piano) believe year (expand mom isomorphic) tribune.

(29) The guest impressed the (remarkably responsible) brother of the (remarkably responsible) groom who (often unknowingly) snores.

x-x-x arise renew plant (tenderize destroy) anymore why hear (tenderize destroy) renew rain (overalls vertebrates) drought.

(30) The supplier liked the (really cheerful) employee of the (really cheerful) gardener who (carefully and dutifully) mowed.

x-x-x misinform frank hear (lose remember) conclude go lung (lose remember) portray neck (elevate fact refresher) datum.

(31) The grandfather bothered the (truly considerate) father of the (truly considerate) boy who (often soundly) slept.

x-x-x constitute hectares grape (hear adventure) because game know (hear adventure) nor got (enjoy scorer) basin.

(32) The client missed the (unusually tall) patron of the (unusually tall) artist who (inconsiderately and insultingly) swore.

x-x-x defund soccer con (despise save) illuminate why she (despise save) ensure lock (villain sir renumbering) cloud.

(33) The neighbor hated the (exceptionally moody) son of the (exceptionally moody) fisherman who (successfully) swam (across the lake).

x-x-x tolerate watts vote (undergo dream) upon drop ago (undergo dream) surpass loss (acknowledge) craze (accept get info).

(34) The cook greeted the (rather short) boss of the (rather short) waitress who (often) drank (after work).

x-x-x rely wording walk (weaken use) aspire nor she (weaken use) believe brush (mile) rider (march lion).

(35) The host entertained the (quite talented) partner of the (quite talented) writer who (sometimes) laughed (at inopportune moments).

x-x-x hear collapse sky (lighten moonshine) depend defy ask (lighten moonshine) ignore him (advantage) vaccine (goal determine expect).

(36) The boss embarrassed the (extremely outgoing) editor of the (extremely outgoing) author who (often) struggled (with the computer).

x-x-x pray cholesterol marry (survive sorority) afford hang else (survive sorority) been leave (him) sanctuary (stay deny tomorrow).

(37) The publicist approached the (genuinely interesting) girlfriend of the (genuinely interesting) actor who (purposely) fasted (before the party).

x-x-x absorb eulogize go (believe sweeten) everywhere move pin (believe sweeten) try want (mythologize) insult (review front yours).

(38) The student saw the (relatively fashionable) colleague of the (relatively fashionable) librarian who (rudely) yawned (during the meeting).

x-x-x anybody feel gem (energize singer) salivate are why (energize singer) bandwidth train (airport) wake (agency guys anymore).

(39) The ranger called the (incredibly talkative) guide of the (incredibly talkative) hiker who (accidentally) fell (down the hill).

x-x-x modify forget see (immerse punishment) compete nor hear (immerse punishment) ignore yes (disable) sale (fund go info).

(40) The reporter avoided the (fairly conservative) intern of the (fairly conservative) politician who (abruptly) departed (from the meeting).

x-x-x moreover flavor yes (grow capitalize) inform leaf wish (grow capitalize) accelerate luck (affirm) capture (done per percent).

(41) The politician met the (really charming) commander of the (really charming) soldier who (courageously) fought (at the battle).

x-x-x unless cup gone (peruse desktop) exaggerate hay have (peruse desktop) anyway nor (discuss) milestone (say zap heard).

(42) The lawyer knew the (highly ethical) clerk of the (highly ethical) judge who (quickly) left (after the hearing).

x-x-x give angry nor (pervade configure) edged me bet (pervade configure) toward fun (generation) lake (dream eat consider).

(43) The neighbor disliked the (exceedingly superficial) friend of the (exceedingly superficial) artist who (frequently) entertained (on the weekends).

x-x-x whether persist go (enchant disparage) afford low stink (enchant disparage) criticize goal (blackboard) saxophone (far guy learn).

(44) The athlete impressed the (particularly handsome) husband of the (particularly handsome) coach who (often) volunteered (at the hospital).

x-x-x disown euthanize we (know disdain) anyway me done (know disdain) accepts hell (show) heraldic (up spy situate).

(45) The manager called the (incredibly attractive) costar of the (incredibly attractive) actor who (loudly) complained (about the rehearsal).

x-x-x neither wash eat (demonstrate duckling) enjoy luck mist (demonstrate duckling) sustain holy (conclusion) defenestration (arm have distinguish).

(46) The bartender watched the (awfully observant) girlfriend of the (awfully observant) nurse who (quietly) laughed (at the joke).

x-x-x encroach whichever be (predict blackboard) assail hell shut (predict blackboard) brighten true (award) movie (soul pin toward).

(47) The teacher met the (genuinely kind) sister of the (genuinely kind) student who (always) sang (in the car).

x-x-x choose us nor (extend push) drank why sure (extend push) thereof goal (definition) yam (shy rug sit).

(48) The counselor helped the (tremendously talented) boyfriend of the (tremendously talented) guitarist who (often) partied (on the weekends).

x-x-x therefore deliver we (investigate balance) behave tall nor (investigate balance) discography swam (photo) downfall (see upon reference).

References

Adams, B. C., Clifton, C., & Mitchell, D. (1991). Lexical guidance in sentence parsing. *Bulletin of the Psychonomic Society*, *29*(6).

Adams, B. C., Clifton, C., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review*, *5*(2), 265–270.

Anderson, C., & Carlson, K. (2010). Syntactic structure guides prosody in temporarily ambiguous sentences. *Language and speech*, *53*(4), 472–493.

Arnett, N., & Wagers, M. (2017). Subject encodings and retrieval interference. *Journal of Memory and Language*, *93*, 22–54.

Ashby, J., & Clifton Jr, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, *96*(3), B89–B100.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, *19*(3), 391–426.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In *Reanalysis in sentence processing* (pp. 1–46). Springer.

Bellik, J. (to appear). Stringency hierarchies in prosodic sisterhood: Strongstart and equalsisters. In *Syntax-prosody in optimality theory: Theory and analyses.* Springer.

Bellik, J., & Kalivoda, N. (2016). Adjunction and branchingness effects in syntax-prosody mapping. In *Proceedings of the annual meetings on phonology* (Vol. 3).

Bellik, J., & Kalivoda, N. (2019). Automated tableau generation using spot (syntax prosody in optimality theory). *Linguistics Vanguard*, *5*(1).

Bellik, J., & Van Handel, N. (to appear). Size effects in prosody: Branch-counting, leaf-counting, and uniformity. In *Syntax-prosody in optimality theory: Theory and analyses.* Springer.

Bennett, R. (2018). Recursive prosodic words in kaqchikel (mayan). *Glossa: A journal of general linguistics*, *3*(1).

Bennett, R., & Elfner, E. (2019). The syntax–prosody interface. *Annual Review of Linguistics*, *5*, 151–171.

Bishop, J. (2020). Exploring the similarity between implicit and explicit prosody: Prosodic phrasing and individual differences. *Language and Speech*, 1–27.

Bishop, J., Chong, A. J., & Jun, S.-A. (2015). Individual differences in prosodic strategies to sentence parsing. In *Icphs.*

Booij, G. (1996). Cliticization as prosodic integration: The case of dutch. *The Linguistic Review*, *13*, 219–242.

Boyce, V., Futrell, R., & Levy, R. (2020). Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, *111*, 104082.

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, *8*(2), 37–50.

Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of memory and language*, *64*(2), 153–170.

Breen, M., & Clifton, C. (2013). Stress matters revisited: A boundary change experiment. *Quarterly Journal of Experimental Psychology*, *66*(10), 1896–1909.

Breen, M., Fitzroy, A. B., & Oraa Ali, M. (2019). Event-related potential evidence of implicit metric structure during silent reading. *Brain sciences*, *9*(8), 192.

Breen, M., Watson, D. G., & Gibson, E. (2011). Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes*, *26*(10), 1532–1562.

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, *45*(1), 58–81.

Carlson, K., Clifton, C., & Frazier, L. (2009). Nonlocal effects of prosodic boundaries. *Memory & cognition*, *37*(7), 1014–1025.

Carlson, K., & Potter, D. (2022). Focus attracts attachment. *Language and Speech*, *65*(2), 491–512.

Carlson, K., & Tyler, J. C. (2018). Accents, not just prosodic boundaries, influence syntactic attachment. *Language and speech*, *61*(2), 246–276.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive psychology*, *42*(4), 368–407.

Clemens, L. E. (2014). *Prosodic noun incorporation and verb-initial syntax* (Unpublished doctoral dissertation). Harvard University.

Clifton, C., Carlson, K., & Frazier, L. (2006). Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic bulletin & review*, *13*(5), 854–861.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, *40*(2), 141–201.

Elfner, E. (2012). *Syntax-prosody interactions in Irish* (PhD dissertation). University of Massachusetts, Amherst.

Elfner, E. (2015). Recursion in prosodic phrasing: Evidence from connemara irish. *Natural Language & Linguistic Theory*, *33*(4), 1169–1208.

Elfner, E. (2018). The syntax-prosody interface: current theoretical approaches and outstanding questions. *Linguistics Vanguard*, *4*(1).

Elordieta, G. (2015). Recursive phonological phrasing in basque. *Phonology*, *32*(1), 49–78.

Féry, C. (2010). Syntax, information structure, embedded prosodic phrasing, and the relational scaling of pitch accents. *The sound patterns of syntax*, 271–290.

Fodor, J. D. (1998). Learning to parse? *Journal of psycholinguistic research*, *27*(2), 285–319.

Fodor, J. D. (2002a). Prosodic disambiguation in silent reading. In *North east linguistics society* (Vol. 32, p. 8).

Fodor, J. D. (2002b). Psycholinguistics cannot escape prosody. In *Speech prosody 2002, international conference.*

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. (1979). *On comprehending sentences: Syntactic parsing strategies.* University of Connecticut.

Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in cognitive sciences*, *10*(6), 244–249.

Frazier, L., Carminati, M. N., Cook, A. E., Majewski, H., & Rayner, K. (2006). Semantic evaluation of syntactic structure: Evidence from eye movements. *Cognition*, *99*(2), B53–B62.

Frazier, L., & Clifton, C. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*(3), 277–295.

Frazier, L., Clifton, C., Carlson, K., & Harris, J. A. (2014). Standing alone with prosodic help. *Language, Cognition and Neuroscience*, *29*(4), 459–469.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology*, *14*(2), 178–210.

Fromont, L. A., Soto-Faraco, S., & Biau, E. (2017). Searching high and low: Prosodic breaks disambiguate relative clauses. *Frontiers in Psychology*, *8*, 96.

Ghini, M. (1993). Phi-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics*, *12*(2), 41–78.

Hemforth, B., Fernandez, S., Clifton Jr, C., Frazier, L., Konieczny, L., & Walter, M. (2015). Relative clause attachment in german, english, spanish and french: Effects of position and length. *Lingua*, *166*, 43–64.

Hirose, Y. (2003). Recycling prosodic boundaries. *Journal of psycholinguistic research*, *32*(2), 167–195.

Hoeks, M., Toosarvandani, M., & Rysling, A. (2021). Decomposing the focus effect: Evidence from reading.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes*, *26*(3), 376–405.

Hofmeister, P., & Vasishth, S. (2014). Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in psychology*, *5*, 1237.

Hwang, H., & Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative np ambiguity in korean. *Journal of Psycholinguistic Research*, *38*(2), 151–175.

Ito, J., & Mester, A. (2009). The extended prosodic word. *Phonological domains: Universals and deviations*, *1*, 135–194.

Ito, J., & Mester, A. (2012). Recursive prosodic phrasing in japanese. *Prosody matters: Essays in honor of Elisabeth Selkirk*, 280–303.

Ito, J., & Mester, A. (2013). Prosodic subcategories in japanese. *Lingua*, *124*, 20–40.

Ito, J., & Mester, A. (2019). Match theory and prosodic well formedness constraints. *Prosodic Studies: Challenges and prospects*, 252–274.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (reply to fitch, hauser, and chomsky). *Cognition*, *97*(2), 211–225.

Jaeger, T. F., Fedorenko, E., & Gibson, E. (2005). Dissociation between production and comprehension complexity. *Poster Presentation at the 18th CUNY Sentence Processing Conference, University of Arizona*.

Jun, S.-A. (2010). The implicit prosody hypothesis and overt prosody in english. *Language and Cognitive Processes*, *25*(7-9), 1201–1233.

Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology*. Academic Press.

Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition*, *123*(1), 1–20.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*(2), 153–194.

Konieczny, L., & Döring, P. (2003). Anticipation of clause-final heads: Evidence from eye-tracking and srns. In *Proceedings of iccs/ascs* (pp. 13–17).

Kroll, M. (2020). *Comprehending ellipsis*. University of California, Santa Cruz.

Ladd, D. R. (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology*, *3*, 311–340.

Lee, E.-K., & Garnsey, S. M. (2012). Do contrastive accents modulate the effect of intonational phrase boundaries in parsing? *Lingua*, *122*(14), 1763–1775.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Levy, R., & Keller, F. (2013). Expectation and locality effects in german verb-final structures. *Journal of memory and language*, *68*(2), 199–222.

Maynell, L. A. (1999). Effect of pitch accent placement on resolving relative clause ambiguity in english. In *Poster presented at the 12th annual cuny conference on human sentence processing, new york* (pp. 18–20).

Myrberg, S. (2013). Sisterhood in prosodic branching. *Phonology*, *30*(1), 73–124.

Nakatani, K., & Gibson, E. (2008). Distinguishing theories of syntactic expectation cost in sentence comprehension: Evidence from japanese. *Linguistics*, *46*(1), 63–87.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Pak, M. (2008). *The postsyntactic derivation and its phonological reflexes* (Unpublished doctoral dissertation). Ph. D. dissertation, U. of Pennsylvania.

Peperkamp, S. A. (1997). *Prosodic words* (Vol. 34). Holland Academic Graphics The Hague.

Prieto, P. (2005). Syntactic and eurhythmic constraints on phrasing decisions in catalan. *Studia Linguistica*, *59*(2-3), 194–222.

Prieto, P. (2006). Phonological phrasing in Spanish. In F. Martínez-Gil & S. Colina (Eds.), *Optimality-Theoretic Studies in Spanish Phonology* (pp. 39–61). Amsterdam: John Benjamins.

Rich, S., Balachandran, L., Duff, J., Kogan, M., Wax Cavallaro, M., Van Handel, N., & Wagers, M. (2022). "it's alive!": Animacy-based structural expectations rapidly update with context. *Short talk at the Conference on Human Sentence Processing (HSP) 35, University of California, Santa Cruz.*

Royer, A. J. (2021). *The role of prosodic phrasing on the acceptability of agreement attraction in english*. University of California, Los Angeles.

Samuels, B. D. (2009). *The structure of phonological theory*. Harvard University.

Sandalo, F., & Truckenbrodt, H. (2002). Some notes on phonological phrasing in brazilian portuguese. *MIT Working Papers in Linguistics*, *42*(1), 285–310.

Sasaki, K. M. (2021). *Components of coherence*. University of California, Santa Cruz.

Schafer, A. J. (1997). *Prosodic parsing: The role of prosody in sentence comprehension*. University of Massachusetts Amherst.

Schafer, A. J., Carter, J., Clifton, C., & Frazier, L. (1996). Focus in relative clause construal. *Language and cognitive processes*, *11*(1-2), 135–164.

Scheer, T. (2012). Chunk definition in phonology: prosodic constituency vs. phase structure. *Modules and interfaces*, 221–253.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999).

Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147.

Seidl, A. (2013). *Minimal indirect reference: a theory of the syntax-phonology interface*. Routledge.

Selkirk, E. (1984). Phonology and syntax: the relation between sound and structure. *Current Studies in Linguistics I*.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology*, *3*, 371–405.

Selkirk, E. (1996). The prosodic structure of function words. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 187–213). Erlbaum.

Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In *Prosody: Theory and experiment* (pp. 231–261). Springer.

Selkirk, E. (2011). The syntax-phonology interface. *The handbook of phonological theory*, *2*, 435–483.

Sloggett, S., Van Handel, N., & Rysling, A. (2020). A-maze by any other name. *Poster presented at the 33rd Annual CUNY Conference on human sentence processing, University of Massachusetts, Amherst*.

Sloggett, S., Van Handel, N., Sasaki, K., Duff, J., Rich, S., Orth, W., … Rysling, A. (2020). “ambiguous” isn’t “underspecified”: Evidence from the maze task.

Poster presented at the 33rd Annual CUNY Conference on human sentence processing, University of Massachusetts, Amherst.

Speer, S. R., & Foltz, A. (2015). The implicit prosody of corrective contrast primes appropriately intonated probes (for some readers). In *Explicit and implicit prosody in sentence processing* (pp. 263–285). Springer.

Staub, A. (2007a). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 550.

Staub, A. (2007b). The return of the repressed: Abandoned parses facilitate syntactic reanalysis. *Journal of memory and language*, *57*(2), 299–323.

Stolterfoht, B., Friederici, A. D., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: Differential erp effects. *Cognition*, *104*(3), 565–590.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216.

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: a psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64.

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.

Van Gompel, R. P., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*(2), 284–307.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In *Reading as a perceptual process* (pp. 621–648). Elsevier.

Vasishth, S., & Drenhaus, H. (2011). Locality in german. *Dialogue & Discourse*, *2*(1), 59–82.

Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 767–794.

Vogel, I. (2009). The status of the clitic group. In J. Grijzenhout & B. Kabak (Eds.), *Phonological domains: Universals and deviations* (pp. 15–46). Berlin: Mouton de Gruyter.

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and cognitive processes*, *25*(7-9), 905–945.

Watson, D., & Gibson, E. (2004a). Making sense of the sense unit condition. *Linguistic Inquiry*, 508–517.

Watson, D., & Gibson, E. (2004b). The relationship between intonational phrasing and syntactic structure in language production. *Language and cognitive processes*, *19*(6), 713–755.

Watson, D., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia linguistica*, *59*(2-3), 279–300.

Webman-Shafran, R., & Fodor, J. D. (2016). Phrase length and prosody in on-line ambiguity resolution. *Journal of psycholinguistic research*, *45*(3), 447–474.

Wilkenfeld, D. C. (1985). *Encoding prosody in silent reading* (Unpublished doctoral dissertation). University of Connecticut.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of psycholinguistic research*, *41*(2), 105–128.