

Inactive gap formation: An ERP study on the processing of extraction from adjunct clauses

Annika Kohrt, Trey Sorensen, Peter O’Neill & Dustin A. Chacón*

Abstract. Filler-gap (movement, extraction, displacement) dependencies are processed *actively*, i.e., comprehenders anticipatorily commit to an interpretation of the sentence before encountering bottom-up evidence. This suggests that comprehenders make structural commitments to how a sentence will unfold shortly after encountering a filler NP. However, the grammaticality of some filler-gap dependencies may depend on semantic and pragmatic features of the sentence that are not typically considered in studies on filler-gap dependencies. One particular case is extraction from adjunct clauses, in which the filler NP may only grammatically be understood as the object of a non-finite adjunct clause if the main verb is an achievement predicate (e.g., *What coffee did you arrive [drinking __]?* Truswell 2011). We present evidence from an EEG study demonstrating that comprehenders do not actively construct filler-gap dependencies in constructions such as these. Instead, they “inactively” build the dependency, only after integrating semantic information about the adjunct clause into the sentence.

Keywords. psycholinguistics; semantics; syntax; neurolinguistics; filler-gap dependencies; islands; EEG

1. Introduction. In a sentence like (1), the phrase *the apple* is understood as the syntactic object and semantic theme of the verb *eat*. To understand such filler-gap dependencies¹, comprehenders must simultaneously process the linguistic material intervening between the filler phrase (*the apple*) and its resolution site while maintaining information about the dependency in working memory. Top-down linguistic knowledge must also be used to identify possible resolution sites, e.g., identifying *ate* as a verb.

(1) This is **the apple** that Ernie ate __ .

Many results demonstrate that the psycholinguistic processes used to understand filler-gap dependencies are *active*, i.e., comprehenders attempt to resolve the filler-gap dependency at the earliest possible position in the sentence (Stowe 1986; Traxler & Pickering 1996). For instance, in sentences like (2), Traxler & Pickering (1996) found increased reading times at the verb *write* compared to controls, indicating that comprehenders had initially constructed an analysis of the sentence in which *the city* was interpreted as the theme of the verb. However, since *write* is not a verb that felicitously accepts *the city* as a theme, the resulting implausible interpretation causes processing difficulty, either due to a subsequent reanalysis process or to

* We would like to thank Jason Overfelt, Brian Reese, Claire Halpert, and Colin Phillips for comments and criticisms, and the audiences of the 31st CUNY Human Sentence Processing and the 54th Chicago Linguistics Society conferences for hearing earlier versions of this work. We’d also like to thank the Center for Advanced Translational Sensory Sciences (CATSS) and the Minnesota Syntax-Psycholinguistics Laboratory at the University of Minnesota, and the Neuroscience of Language Lab at NYU Abu Dhabi. This work was supported by an Undergraduate Research Opportunity Program (UROP) award to Annika Kohrt. Annika Kohrt, University of Minnesota (kohrt008@umn.edu), Trey Sorensen, University of Minnesota (soren870@umn.edu), Peter O’Neill (petejmoneill@gmail.com), Dustin A. Chacón, New York University: Abu Dhabi (dustin.alfonso@nyu.edu)

¹ Also known as displacement, movement, or extraction dependencies. We will use the term “filler-gap dependency” when highlighting the aspects of relating a filler phrase and a later position over time, and “extraction” when highlighting representational aspects of the dependency.

the surprisal of encountering this word. This increased processing cost is called *the plausibility mismatch effect*.

(2) This is **the city** that the author wrote enthusiastically about -- .

Although active dependency formation processes are well-demonstrated, the precise formulation of these processes has not been carefully described. For instance, they can be cast as a prospective search for a resolution site that is launched shortly after encountering the filler phrase, a re-ranking of preferred parses of the sentence (Frazier 1987; Frazier & Clifton Jr. 1989), or an anticipatory predictive process of upcoming syntactic/semantic material (Omaki et al. 2015; Aoshima et al. 2004; Chacón 2019). On this latter interpretation, the reason the plausibility mismatch effect is observed in sentences like (2) is because, after encountering *the city*, comprehenders build a sufficiently detailed analysis of the sentence in which *the city* is interpreted as the theme of the verb and binds a projected gap in the direct object position, (3). After encountering the verb *wrote* and incorporating it into this previously-constructed analysis, comprehenders detect the implausible interpretation. Importantly, the plausibility mismatch effect demonstrates that comprehenders make a (syntactic and semantic) structural commitment *before* encountering the resolution site, i.e., the plausibility of the interpretation of the dependency in the context of the sentence does not appear to assert an effect early in the processing of the dependency (see also Omaki et al. 2015). Indeed, theories of active dependency formation privilege structural properties over, e.g., probability, plausibility, or predictability (Pritchett 1991; de Vincenzi 1991; Aoshima et al. 2004).

(3) This is **the city** that...

Syntactic analysis: [_{CP} This is [the city]_i [_{CP} OP_i that [_{TP} subject [_{VP} verb --_i]]]]

Semantic analysis: $\exists x,e$ [city(x) & theme(e,x) & verb(e)]

Filler-gap dependency processes are active, but they are also selective. Although filler-gap dependencies may cross indefinitely many words, (4-a), there are a set of grammatical configurations that they may not cross without a significant reduction in acceptability, referred to as syntactic islands (Ross 1967), (4-b). These phenomena are typically attributed to well-formedness conditions on the representation of the sentence (Ross 1967; Chomsky 1977, 1981; Huang 1982; Uriagereka 1999; Rizzi 1990, 2013, but see Deane 1991; Kluender & Kutas 1993; Kluender 1998; Hofmeister & Sag 2010). Importantly, evidence of active dependency formation for filler-gap dependencies is not observed in island configurations. That is, comprehenders only actively consider filler-gap dependencies that are also perceived to be acceptable. This finding has been interpreted as reflecting a tight coupling between grammatical constraints and processing strategies (Phillips 2006, 2013; Yoshida et al. 2014; Omaki et al. 2015). If filler-gap dependency processing depends on fine-grained structural prediction mechanisms, as we envision, then such island-sensitivity can be characterized as a quick and accurate use of structural information to distinguish possible vs. impossible predicted structures.

(4) a. This is **the apple** that Jodie knows ... [that Ernie ate --].

b. *This is **the apple** that Jodie wonders [whether Ernie ate --].

Although island phenomena are typically cast as syntactic constraints, (compositional and conceptual) semantic and pragmatic factors are observed to affect the patterns of acceptability (Goldsmith 1985; Lakoff 1986; Kehler 2002; Truswell 2007a, 2011; Szabolcsi 2006; Ambridge

& Goldberg 2008). Here, we specifically focus on the status of adjunct clauses. Traditionally, tensed adjunct clauses are thought to be syntactic islands (Huang 1982; Chomsky 1986; Cinque 1990; Uriagereka 1999), due to facts like in (5).

- (5) a. This is **the apple** that Ernie said [that he wanted to eat --]
 b. ?*This is **the apple** that Ernie cried [because he wanted to eat --]
 c. ?*This is **the apple** that Ernie slept [before he ate --]
 d. ?*This is **the apple** that Ernie smiled [while he ate --]

However, Robert Truswell has identified a number of semantic factors that modulate the acceptability of extraction from adjunct clauses (Truswell 2007a,b, 2011). For instance, extraction from a tensed adjunct clause is perceived to be worse than extraction from an untensed one, (6). Among untensed adjunct clauses, extraction from a predicate denoting an activity (in the sense of Vendler 1967; *eating the apples*) is more acceptable when the main predicate is an achievement (e.g., *arrive*), rather than another activity predicate (e.g., *work*), as in (7).

- (6) a. ?These are **the apples** that Ernie smiled [while eating --]
 b. *These are **the apples** that Ernie smiled [while he ate --]
- (7) a. ?These are **the apples** that Ernie arrived [eating --]
 b. ?*These are **the apples** that Ernie worked [eating --]

To account for such contrasts, Truswell proposed that extraction from an adjunct clause is grammatical if it's possible to construe the main and adjunct predicates as codescribing the same event. Formally, this may occur if the event variable introduced by the adjunct clause can be identified with the event variable introduced in the main clause (Kratzer 1996). For instance, extraction from tensed adjuncts, (6-b), is perceived to be worse than untensed adjuncts, because the event variable introduced by the adjunct clause is bound by the existential quantifier introduced with tense, (8), thus preventing event identification. More crucially for our study, the contrast in (7) is explained by appealing to the representation of achievement predicates. Following Higginbotham (2009), Truswell proposes that achievement predicates² contain an unbound event variable, corresponding to the events or state of affairs preceding the event of culmination lexicalized by the predicate, (9-a). Thus, the event variable introduced by a non-finite adjunct clause containing an activity predicate can be identified with this extra main clause event variable, yielding a macro-event, represented by E in (9-b). This mode of composition is unavailable when the main verb is an activity predicate, (9-c).

- (8) $\llbracket \text{the apples that Ernie smiled while he ate} \rrbracket =$
 $\lambda y. \exists e[\text{smile}(e) \ \& \ \text{agent}(e, \text{ernie})] \ \& \ \exists e'[\text{eat}(e') \ \& \ \text{theme}(e', y) \ \& \ \text{apples}(y)]$

- (9) a. $\llbracket \text{arrive} \rrbracket = \lambda e. \text{arrive}(e) \ \& \ E[\langle e', e \rangle]$

² Truswell also distinguishes between *true achievements* and *points*. True achievements lexicalize a culmination point, and imply a previous process (represented as the unbound event variable), whereas *point* lack this previous process. This can be seen in (i), adapted from (Truswell 2007b, 2011) in which the true achievement *reach the summit* denotes the pre-culmination process when put in the progressive aspect, whereas the point *noticing* does not. The analysis sketched here only applies to true achievements.

- (i) a. I'm reaching the summit as we speak
 b. I'm noticing the problem as we speak

- b. $\llbracket \text{arrived eating } _ _ \rrbracket = \lambda x. \exists e, e' [\text{arrive}(e) \ \& \ E[\langle e, e' \rangle] \ \& \ \text{eat}(e') \ \& \ \text{theme}(e', x)]$
- c. $\llbracket \text{worked eating } _ _ \rrbracket = \lambda x. \exists e, e' [\text{work}(e) \ \& \ \text{eat}(e') \ \& \ \text{theme}(e', x)]$

Although the form of explanation for the contrasts above focuses on the nature of the compositional semantics of the main and adjunct predicate, Truswell (2011) is cautious to observe that aspects of pragmatics and world knowledge are relevant for determining the island status of an adjunct clause as well. For instance, (10-b) is less acceptable than (10-a), because the event of dripping mud on a carpet is unlikely to precede an event of entering a home, given the world knowledge that carpets are usually indoors and not outdoors. Thus, aspects of the compositional semantics (tense, lexical aspect) appear to interact with world knowledge in determining whether a given filler-gap dependency is perceived to be acceptable. If these generalizations are robust, then this suggests that the island status of adjunct clauses is affected by factors that are not typically considered in theories of filler-gap dependency processing.

- (10) a. This is **the carpet** that John exited the house [dripping mud on $_ _$]
 b. This is **the carpet** that John entered the house [dripping mud on $_ _$]

In our work reviewed in the next section (Kohrt et al. 2018, 2019), we have sought to characterize the psycholinguistic processes involved in understanding filler-gap dependencies crossing into adjunct clauses. Specifically, we were interested in seeing whether comprehenders were capable of leveraging aspects of the semantics of the main predicate to selectively project a gap in an adjunct clause. We wanted to know whether, before hearing the verb *drinking* in a sentence like (11), comprehenders actively considered an interpretation in which *the coffee* would then be interpreted as its object.

- (11) John prepared **the coffee** that his best friend arrived at the office [_{Adjunct CP} drinking $_ _$]

For sentences of this type, we failed to find the typical processing profile associated with filler-gap dependencies. Instead, we found a novel pattern of results³, in which processing time increased for plausible filler-gap dependencies in configurations, i.e., there was an additional processing cost associated with a plausible interpretation, rather than an implausible one. We argued that this reflected bottom-up, “inactive” dependency completion, which we attributed to the unusual status of extraction from adjunct clauses generally. In this paper, we strengthen the empirical basis for this novel pattern of results by replicating this effect using electroencephalography (EEG), and relate it to the broader question of what kind of parsing strategies underpin filler-gap dependency processing.

2. Previous findings. In previous work, we examined whether the reported contrast in (7) is reflected in real-time sentence processing. In contrast to much of the work on the processing of filler-gap dependencies characterizes the processes in syntactic terms, we sought to determine whether comprehenders were capable of using semantic features of the sentence to selec-

³ Colin Phillips (p.c.) points out that an increase in processing time associated with a plausible interpretation has been found for other dependencies. For instance, Kazanina & Phillips (2010) found that reading times increased in processing a cataphoric dependency (“backwards anaphora”, a coreference relation in which the pronoun precedes the antecedent) that violates a grammatical rule of Russian. They argued that this pattern of results indicates that comprehenders initially built the illicit interpretation, and then ruled it out using grammatical knowledge as a filter.

tively project a gap in an adjunct clause. One hypothesis that we considered was that comprehenders pursued a permissive strategy, and considered every non-finite adjunct clause as a possible host for a gap. This predicted evidence of active dependency formation in all non-finite adjuncts, regardless of the identity of the main verb. Alternatively, we hypothesized that comprehenders could selectively apply active gap formation in non-finite adjunct clauses depending on the semantic aspects of the previous verb phrase. If so, we predicted evidence of active gap formation in non-finite adjunct clauses, but only if the previous verb was an achievement. This analysis would imply that comprehenders could quickly incorporate some semantic information into the analysis, and use this information to revise a predicted structure.

To test this, we conducted two self-paced reading experiments (reported in Kohrt et al. 2018 and Kohrt et al. 2019), using a plausibility mismatch paradigm, exemplified in (12). We expected to find increased reading times associated with the implausible filler-gap dependencies (*the report ... drinking*) compared to the +Plausible controls. We did not find evidence of a plausibility mismatch effect in the comparison of the levels of \pm Plausibility within the levels of \pm Composable. Instead, we found that processing time increased when the filler was a plausible argument, i.e., a “reverse” plausibility mismatch effect.

- (12) a. +PLAUSIBLE, +COMPOSABLE
 John prepared **the coffee** that his best friend **arrived** at the office [Adjunct CP drinking yesterday afternoon]
- b. +PLAUSIBLE, –COMPOSABLE
 John prepared **the coffee** that his best friend **worked** at the office [Adjunct CP drinking yesterday afternoon]
- c. –PLAUSIBLE, +COMPOSABLE
 John prepared **the report** that his best friend **arrived** at the office [Adjunct CP drinking yesterday afternoon]
- d. –PLAUSIBLE, –COMPOSABLE
 John prepared **the report** that his best friend **worked** at the office [Adjunct CP drinking yesterday afternoon]

As a *post hoc* interpretation of this result, we suggested that comprehenders do not initially attempt to build a filler-gap dependency into an adjunct clause. Instead, the semantic information of the filler phrase is maintained while the adjunct clause is integrated into the syntactic and semantic analysis of the sentence, (13-a). In the case of processing an achievement predicate, integrating the adjunct clause into the global analysis of the sentence then cues the construction of a macro-event, which may also require accessing information about world knowledge from memory, (13-b). If the filler NP is a suitable theme of the event described by the adjunct event and this resulting macro-event, then comprehenders construct a gap in a “bottom-up” manner, given that the main predicate-adjunct predicate complex codescribes a single event, (13-c). This extra step in accommodating an otherwise unlicensed gap triggers increased processing time.

- (13) a. **Step 1:** Maintain filler phrase, process adjunct clause:
Syntax: [the coffee_i ... [VP [arrive] [drinking]]
Semantics: ... arrive(e) & drink(e') & coffee(x)
- b. **Step 2:** Relate events:

Syntax: [the coffee_i ... [VP [arrive] [drinking]]
Semantics: ... arrive(e) & drink(e') & coffee(x) & E[⟨e,e'⟩]

c. **Step 3: Bottom-up gap identification:**

Syntax: [the coffee_i ... [VP [arrive] [drinking --_i]]
Semantics: ... arrive(e) & drink(e') & coffee(x) & E[⟨e,e'⟩] & theme(e',x)

This contrasts with our model of typical filler-gap dependency processing in several ways. First, instead of anticipatorily building a syntactic and semantic analysis to accommodate the resolution of the filler-gap dependency, comprehenders construct a resolution site for the dependency only after processing some of the adjunct clause. Additionally, instead of using the presence of the filler-gap dependency to project a partial semantic structure that may result in an implausibility, we propose that comprehenders incorporate plausibility information first, while relating the (sub)vents described by the main clause and the adjunct clause, and then selectively resolve the dependency. In other words, whereas typical filler-gap dependency involve early commitment to specific structures which then can be either quickly confirmed or disconfirmed, processing filler-gap dependencies resolving into an adjunct clause is a slower process that is gated initially by building a semantic (and conceptual) representation.

However, the fine-grained interpretation that we gave these results is underdetermined by the results themselves, e.g., there is no distinct signal in the data indexing initial surprisal that the verb does not match the initially predicted structure vs. the hypothetical later bottom-up gap completion process we described above. For this reason, we opted to use EEG, due to its increased temporal latency and its increased dimensionality which allows for distinguishing different phases of processing. Previous EEG studies on the processing of filler-gap dependencies show electrophysiological activity indexing the completion of the filler-gap dependency within several hundred milliseconds of encountering the verb, described in the next section. However, we show that the event-related potential (ERP) record for sentences like (12) do not diverge until significantly later than this, consistent with the inactive, meaning-driven, bottom-up parsing strategy that we sketched above.

3. Experiment.

3.1. RATIONALE. The goal of the experiment is to determine whether comprehenders actively build filler-gap dependencies into adjunct clauses. Previous self-paced reading results failed to demonstrate evidence of active dependency formation, but instead yielded an increase in processing time for plausible sentences, which we interpreted as evidence of a bottom-up mode of composition. However, because self-paced reading data depends on a motor response that may not be closely time-locked with comprehension processes, it does not give a high-resolution snapshot of the distinct phases of processing. Thus, we chose to conduct an EEG study, because this technique has higher temporal resolution. Plus, EEG provides different indices of processing (“ERP components”), which provide a more detailed picture of the underlying mechanisms.

Much of the discussion above has focused on the initial detection of the plausibility of a filler-gap dependency as detected in reading times. However, Garnsey et al. (1989) demonstrated a similar sensitivity to (im)plausibility of a filler-gap dependency in EEG. They found that, at the verb *called*, the amplitude of the N400, a negative-going deflection that peaks around 400ms after onset of stimulus, was larger in sentences like (14-a) than in (14-b), which they

	N400	P600	Later
Active-gap formation	Effect of –Plausible	Effect of +Composable	
Inactive-gap formation			Effect of +Plausible, +Composable

Table 1. Predicted results on N400 and P600 component.

attributed to the difference in plausibility of the filler-gap dependency at this point in the sentence. The N400 component has traditionally been understood to index difficulty in incorporating a word or phrase into the semantic representation of the sentence (Kutas & Federmeier 2011). However, this interpretation remains controversial, e.g., amplitude of the N400 component may instead index difficulty in accessing a lexical item, with sentence-level predictions facilitating or inhibiting these processes (Lau et al. 2008). Regardless of the precise interpretation of this component, this result demonstrates that in typical active dependency formation situations, comprehenders have incorporated some information about the relation between the filler phrase and the verb before 200-400ms post-onset of the verb.

- (14) a. The businessman knew **which article** the secretary **called** at home
b. The businessman knew **which customer** the secretary **called** at home

Another component that has been identified as reflecting filler-gap dependency processes is the P600, a positive deflection that typically begins around 600ms after onset of the stimulus, which is observed in many contexts that trigger processing difficulty (King & Kutas 1995; Kaan et al. 2000; Phillips et al. 2005; Gouvea et al. 2009). For instance, the P600 is observed upon completion of a filler-gap dependency, and the latency of the component is affected by the length of the dependency (Phillips et al. 2005). Thus, differences in the latency and amplitude of the P600 component may reflect differences in whether or when a filler phrase is integrated into the gap site.

If comprehenders are capable of actively completing a filler-gap dependency into an adjunct clause, then we predict that the amplitude of the N400 component should increase for implausible filler-gap dependencies over plausible ones, replicating the early sensitivity to plausibility reported by Garnsey et al. (1989). Similarly, if comprehenders only selectively build filler-gap dependencies into adjuncts if the previous verb was an achievement predicate, then we predicted an effect of \pm Composability on the P600 component. By contrast, if comprehenders complete filler-gap dependencies in the bottom-up fashion sketched in (13), then we predict that the ERP signal of our sentences should be similar until significantly later, e.g., after 800-1000ms. That is, if the psycholinguistic processes involved in resolving a filler-gap dependency into an adjunct clause are dependent on integrating the meaning of the adjunct clause into the analysis first, then we predict that the effect should follow the traditional language-related ERP components associated with processing the adjunct clause verb. These predictions are sketched in Table 1.

3.2. MATERIALS. We prepared 144 sets of items, with 4 items per set, using the same design as presented in (12). Each item consisted of a preamble (*John prepared*), followed by a filler NP that was either a plausible theme of the verb located in the adjunct clause (*the coffee*,

+Plausible) or not (*the report*, –Plausible). Following this, there was a relative clause pronoun (*that*), followed by a 3-word subject NP (*his best friend*), then either an achievement verb (*arrived*, +Composable) or an activity or accomplishment verb (*worked*, –Composable). Afterwards, there was a three-word long PP (*at the office*). We tried to keep these PPs the same across conditions within each item set, but this was not possible in all items due to the argument structures of the previous verb. However, the final word in this PP was the same across all conditions. Afterwards, the critical region was the non-finite adjunct clause verb (*drinking*), following by a 2 or 3 word spillover region (*late this afternoon*), which were identical across conditions within each item set.

The verbs that we selected for the main verbs and for the adjunct verbs were controlled, such that the +Composable verbs were always paired with the same –Composable verbs across item sets. The pairs of main verbs were *arrive/work*, *reach/wait*, *leave/chat*, *enter/play*, *exit/ride*, and *go/watch*. This meant that there were 24 sets of items associated with each pair of verbs. We also sought to make the identify of the adjunct verb less predictable by counterbalancing which non-finite adjunct verb was paired with which main verbs. There were six verbs we used for this, *drinking*, *reading*, *singing*, *drawing*, *eating*, *knitting*. These verbs were selected because they are equally interpretable as a transitive verb (i.e., with the filler NP as its direct object) or as intransitive. This design choice was made to avoid forcing participants to interpret the filler-gap dependency as resolving with this verb solely due to its transitivity requirements. Counter-balancing these features across all items resulted in 5 sets of items per pair of adjunct verbs and main verb pair.

We also included 256 complexity-matched fillers, with 50% containing minor grammatical violations, semantic anomalies, or both. The items were distributed in a 2×2 Latin Square design, such that each participant saw one item from each set of items, and all the filler items.

3.3. METHODS. Participants ($N = 24$) from the University of Minnesota community sat in a dimly lit, sound-proof room in front of a computer at the Center for Advanced Sensory Sciences (CATSS). We affixed a 32 Ag/Cl electrode BioSemi cap to the participant’s scalp, and applied conductive gel to keep each electrode’s resistance below 40 k Ω . Afterwards, participants were instructed to silently read sentences displayed on the screen. The sentences were displayed in a rapid serial visual presentation (RSVP) design, in which each word automatically displayed one at a time for a fixed duration of 500ms, with an 83ms inter-word interval. The last word of each sentence lasted for 900ms. Electrical potentials were recorded in real-time, referenced to two electrodes affixed to the mastoids. Twenty percent of trials were followed by a comprehension, and participants responded using the keyboard to provide a response.

Participants were given feedback on the screen if they made an incorrect answer. After the other 80% of trials, participants were instructed to press a key to advance to the next sentence. Each sentence was preceded by a fixation cross which also appeared for 500ms. Sentences were displayed in a random order, in 4 distinct blocks. From start to finish, the task took approximately 90 minutes. Participants were allowed to take breaks between each block, and they communicated with the researcher through a baby monitor. Participants were compensated 20USD per hour for their time.

3.4. RESULTS. Before analysis, we applied a 1Hz high-pass FIR filter to the raw EEG signals. Then, we extracted epochs at each verb in the target sentences. The epochs ranged from

100ms before onset of the adjunct verb (*drinking*) to 1600ms afterwards, which contains the next two words in the spillover region. The epochs were baselined to the average potential 100ms before onset of the adjunct verb. This word was identical within each item set across conditions. Then, we automatically rejected epochs that exceeded a peak-to-peak threshold of 150 mV. Three subjects had a large artefact rejection rate, and thus were excluded from analysis. Two additional subjects were removed due to poor recording quality. After epoch rejection, we then applied ICA correction to remove eye-blinks and other periodic artefacts.

We began by analyzing the traditional language-related ERP components. First, we averaged each subject's voltage at each electrode within two critical time windows: the N400 time window (300–500ms) and the P600 time window (600–800ms). Then, we divided the electrodes into $2 \times 2 + 1$ regions – left posterior (P7, P3, PO3, O1), right posterior (P8, P4, PO4, O2), left anterior (F7, F3, AF3, Fp1), right anterior (F8, F4, AF4, Fp2), and midline (Fz, Cz, Pz, Oz), excluding other electrodes. We then conducted two linear mixed effects models using the lmer package in R (Bates et al. 2015; R Core Team 2017). The first mixed effects model that we conducted had a $2 \times 2 \times 2 \times 2$ structure, with Laterality (left, right), Anteriority (anterior, posterior), \pm Plausibility, and \pm Composability as factors, and with participant as random effect in these two time windows. Then, in the midline electrodes, we conducted a 2×2 linear mixed effects model with \pm Plausibility and \pm Composability as factors and with participant as a random effect.

In the N400 time window, the first mixed effects model indicated that voltage was more negative over the entire scalp for +Plausible filler-gap dependencies ($\beta = -0.80$, SE= 0.39, $t = -2.0$, $p = 0.04$), and +Plausible, –Composable sentences had more negative voltage in posterior regions ($\beta = -1.59$, SE= 0.79, $t = -2.0$, $p = 0.04$). However, no post-hoc pairwise comparison conducted on this model within any of the factors was significant ($p > 0.05$). Similarly, the mixed effects model fit to the midline electrodes had no significant intercept in this time window ($p > 0.05$). Thus, although there were significant intercepts in the first model, it is difficult to interpret this as a traditional N400 effect. This is because voltage was more negative for plausible filler-gap dependencies, whereas typically N400 amplitude is more negative for words that are implausible, surprising, or unpredictable given the context. Additionally, the N400 effect is typically observed over midline electrodes, but the significant interaction intercept in the model suggested that the effect was driven by posterior electrodes, which is an atypical scalp distribution for an N400 effect. The lack of traditional N400 effect is underscored by the lack of significant effect in the midline electrode model, where the effect would be expected.

In the P600 time window, there were no significant intercepts in the first model ($p > 0.05$). However, the midline model showed that voltage was more negative for –Composable conditions ($\beta = -1.33$, SE= 0.47, $t = -2.9$, $p < 0.01$). Pairwise comparisons revealed that voltage was more positive for +Composable conditions over –Composable conditions within both levels of \pm Plausibility ($p < 0.05$), suggesting that this is an effect of the combination of the two verb phrases, and thus does not necessarily reflect any interaction between the plausibility of the filler-gap dependency and the identity of the previous verb phrases. Thus, this apparent P600 effect may reflect processes relevant to syntactically or semantically combining the adjunct clause to the preceding clause, not necessarily the integration of the filler NP into the analysis.

Afterwards, we conducted a cluster-based permutation test. This is a non-parametric test

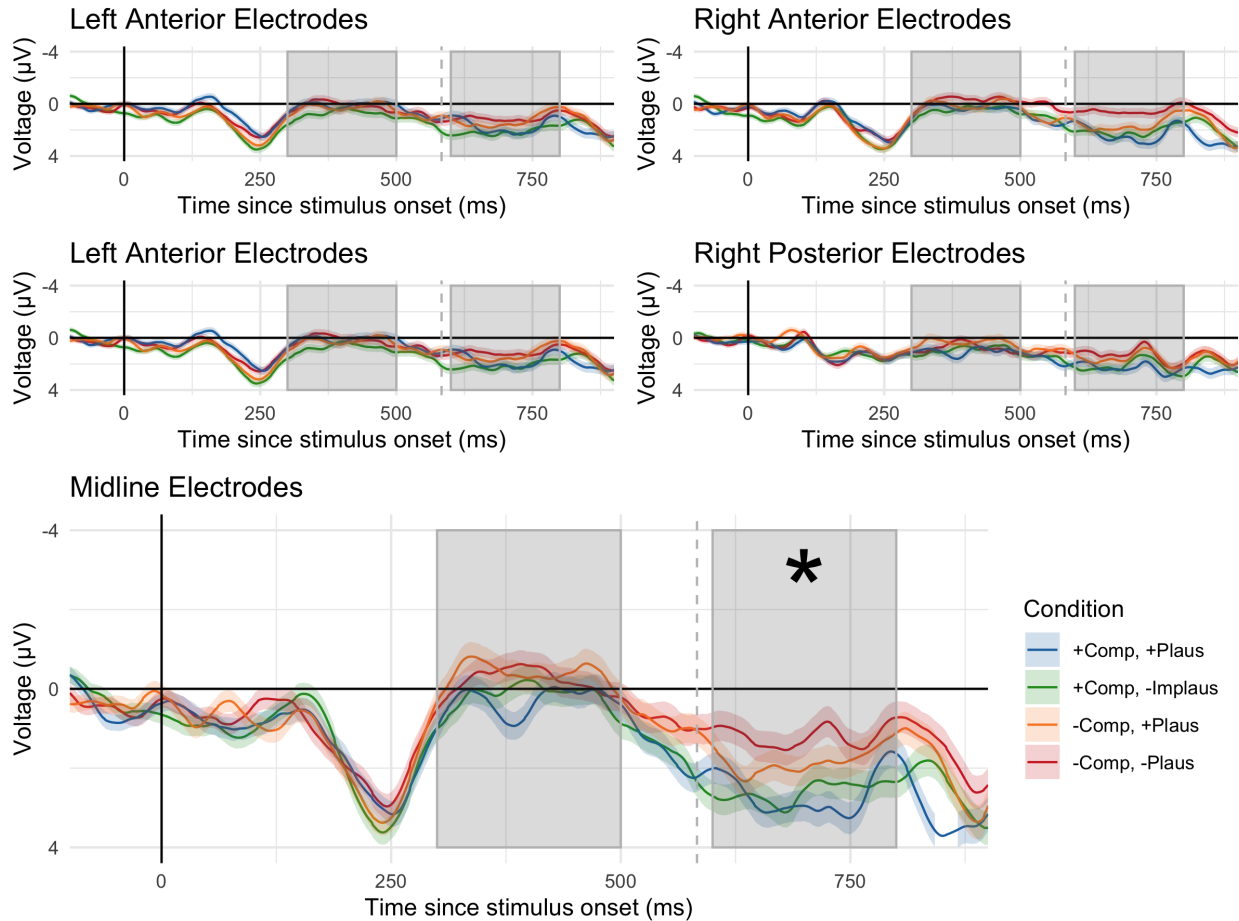


Figure 1. Average voltage by region across conditions. Error ribbons correspond to two standard errors from the mean. Grey areas correspond to N400 time-window (300–500ms) and P600 time-window (600–800ms). Stars indicate that pairwise comparisons conducted on linear mixed effects model revealed a significant effect in this time window. Voltages were low-pass filtered at 20Hz for graphing purposes only.

that seeks to find contiguous points in time and space (electrodes) that are exchangeable, then conducts the significance testing within this cluster (Maris & Oostenveld 2007). This technique is useful when there is no specific spatial or temporal region that we expect to find an effect in *a priori*. We conducted cluster-based permutation tests in 4 distinct time windows from the onset of the verb, 500ms length each, excluding the first 10ms from onset of the critical verb to and the last 10ms of the final time window. The permutation test had a minimum time of 10ms for the clusters, an α -level of 0.05, and used 5000 samples. We found a cluster that ranged from 1236–1309ms over 21 electrodes. Spatially, this cluster was distributed over most of the scalp, but had the strongest effect in the posterior left quadrant, around electrode CP5. Within this cluster, there was a marginally significant interaction between \pm Plausibility and \pm Composability ($p = 0.09$). Although this cluster was only 70ms long, visual inspection revealed that many electrodes appeared to demonstrate a sustained positivity for +Plausible, +Composable conditions throughout this time window, which may peak in the left posterior

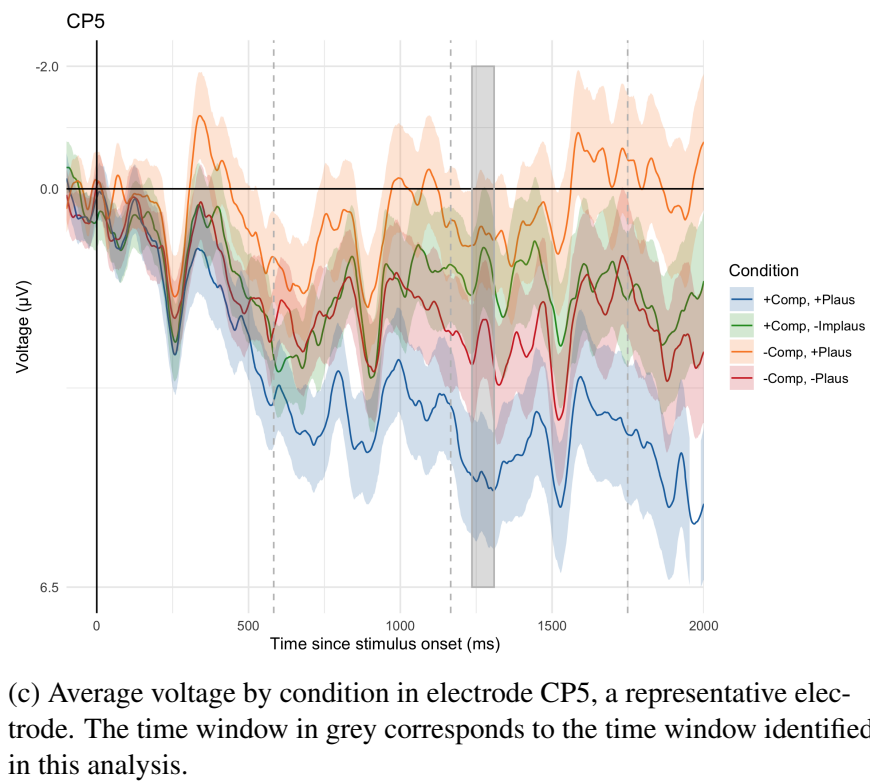
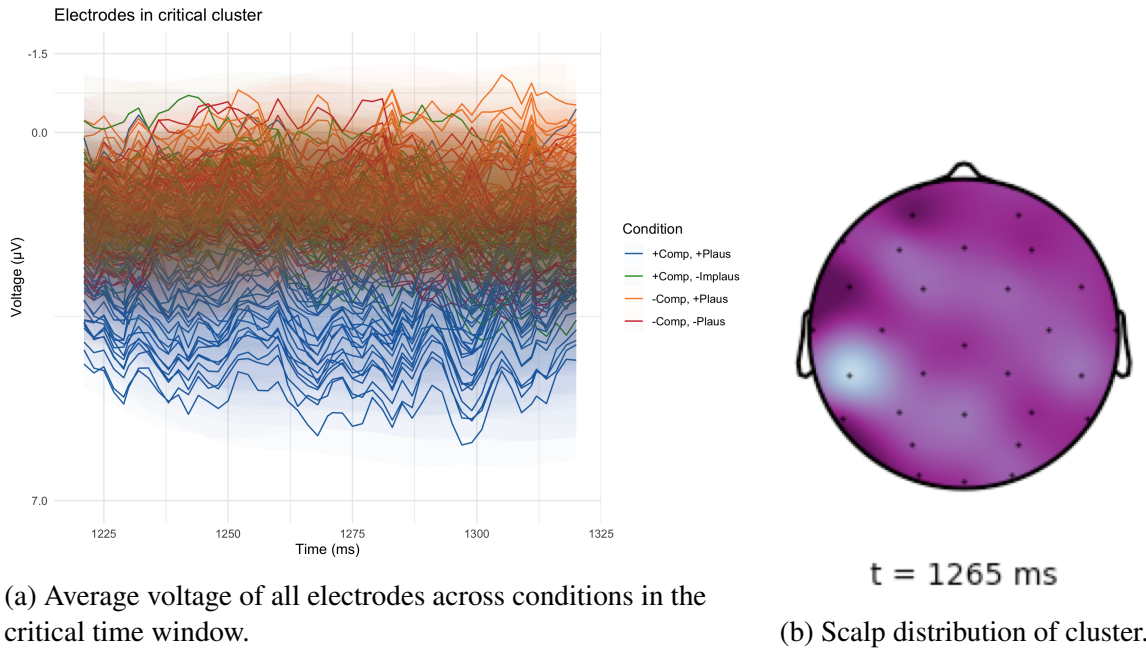


Figure 2. Voltage by condition at time window identified by cluster-based permutation analysis, 1236–1309 ms.

region. This is graphed in Figure 2. Thus, it appears that the effect of processing a plausible filler-gap dependency into an adjunct clause is sensitive to the choice of previous verb, and that these processes are detectable more than two words after the onset of the verb, consistent

with an analysis in which comprehenders construct the dependency in a “bottom-up” fashion, as we argued previously.

4. General discussion and conclusion. To understand a sentence like (11), we argued that comprehenders do not deploy the typical active filler-gap dependency processing strategies. Instead, comprehenders use an “inactive” strategy, in which the syntactic and semantic characteristics of the adjunct clause are first computed, and then the filler-gap dependency is resolved on the basis of these features. We argued for this processing strategy previously on the basis of behavioral data (Kohrt et al. 2018, 2019), however, the low resolution of the self-paced reading technique did not provide a clear enough snapshot to support this analysis. In this paper, we used EEG, which provides a clearer snapshot into these processes, due to its higher temporal resolution and the higher dimensionality of the EEG data. Consistent with our proposal, we failed to find any evidence of active filler-gap dependency formation in the typical language-related ERP components, the N400 and P600. Instead, the effect of extraction from an adjunct clause was observed significantly later, more than 1200ms after the onset of the critical verb.

Despite this, we remain cautious in our interpretation of the results, because there are some methodological and theoretical concerns related to the design of this and our previous studies. First, the experiment presented in this paper has a low sample size. Although we conducted the study with 24 participants, we were only able to report on 19, due to data hygiene issues. This means that higher-powered studies will be required in the future to further strengthen our claims.

Secondly, in all four of our target conditions, the filler-gap dependency had no grammatical resolution. Thus, comprehenders may have been forced to interpret the filler phrase as an argument of the adjunct verb regardless. Because of this design issue, we have no condition that independent tests for the effect of integrating the adjunct clause with the main predicate, separate from interpreting the filler-gap dependency. That is, we have no measure of the brain activity of associating a main clause achievement predicate with an activity adjunct clause vs. two activity predicates. Future studies will need to incorporate a condition with no adjunct clause to more directly isolate this effect.

Similarly, although we attempted to control for the transitivity of the critical adjunct verb (*drinking*), we did not carefully consider the effect of transitivity of the previous verb (*arrived/worked*). This is perhaps an unavoidable confound, since argument structure is heavily correlated with lexical aspect (e.g., Levin 1993; Borer 2005). It is possible that there are differences in processing strategies deployed due to these syntactic differences in the previous clause that do not specifically reference the semantic/conceptual aspects of the clause. Strengthening our claims would require more carefully testing this hypothesis, or finding a new way to remove this confound.

Moreover, the design in this experiment was intended to maximally mirror the design of our self-paced reading experiments, in which we strived to have sufficient spill-over regions between the critical words. However, this introduced a number of issues. First of all, it increased the length between the filler-phrase and intended resolution site. Increased dependency length is shown to diminish the sensitivity to plausible filler-gap dependencies, i.e., comprehenders show a later or smaller plausibility mismatch effect for longer dependencies than shorter dependencies (Wagers & Phillips 2014; Chow & Zhou 2019). Similarly, because our filler-gap dependencies were relative clauses that modified an object of a higher clause

(John prepared *the coffee* that... drinking late this afternoon), this introduces an attachment ambiguity that may affect the EEG results. The non-finite adjunct clause may modify either embedded clause (*arrived/worked*) or the main clause (*prepared*). Although we assume something like the Late Closure strategy should strongly bias participants to attach the adjunct clause to the embedded clause (Frazier & Fodor 1978), the availability of the alternate parse may contribute additional processing difficulty. This is critical to the issue of filler-gap dependency resolution, since no c-command relation holds between the filler NP and the adjunct clause if it is attached at the main clause level, which is required if comprehenders are to consider the adjunct clause as a potential resolution site for the filler-gap dependency. On-going follow-up work eliminates these confounds by using a much simpler design.

Finally, a word of caution about the statistical analyses of our results. Perhaps the most compelling result is the lack of effect of our factorial design on the N400 and the P600 components. However, as always, null results should be treated skeptically, particularly with low sample sizes as reported in the current experiment. Similarly, as Sassenhagen & Draschkow (2016) point out, the cluster-based permutation test is not a significance test of the spatial or temporal location of an observed effect. The cluster-formation phase of the test is only used to identify clusters for computing the test statistic. In other words, the timing and scalp distribution of the effect may be different in possible follow-up experiments. Thus, these findings should only be taken as a starting point for the study of how semantic information interferes in active dependency formation processes.

References

- Ambridge, Benjamin & Adele Goldberg. 2008. The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19(3). 349–381. <https://doi.org/10.1515/COGL.2008.014>.
- Aoshima, Sachiko, Colin Phillips & Amy Weinberg. 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51(1). 23–54. <https://doi.org/10.1016/j.jml.2004.03.001>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Borer, Hagit. 2005. *Structuring sense volume 2: The normal course of events*. Oxford: Oxford University Press.
- Chacón, Dustin A. 2019. Minding the gap?: Mechanisms underlying resumption in English. *Glossa: A Journal of General Linguistics* 4(1). 68. <https://doi.org/10.5334/gjgl.839>.
- Chomsky, Noam. 1977. On wh-movement. In Peter Culicover, Adrian Akmajian & Thomas Wasow (eds.), *Formal syntax*. 71–133. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chow, Wing-Yee & Yangzi Zhou. 2019. Eye-tracking evidence for active gap-filling regardless of dependency length. *The Quarterly Journal of Experimental Psychology* 72(6). 1297–1307. <https://doi.org/10.1177/1747021818804988>.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- Deane, Paul. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2. 1–63. <https://doi.org/10.1515/cogl.1991.2.1.1>.
- Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic*

- Theory* 5. 519–559. <https://doi.org/10.1007/BF00138988>.
- Frazier, Lyn & Charles Clifton Jr. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4. 93–126. <https://doi.org/10.1080/01690968908406359>.
- Frazier, Lyn & Janet Dean Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6(4). 291–325. [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1).
- Garnsey, Susan M., Michael K. Tanenhaus & Robert M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18. 51–60. <https://doi.org/10.1007/BF01069046>.
- Goldsmith, John A. 1985. A principled exception to the coordinate structure constraint. *Proceedings of the Chicago Linguistics Society (CLS)* 21. 133–143.
- Gouvea, Ana, Colin Phillips, Nina Kazanina & David Poeppel. 2009. The linguistic processes underlying the P600. *Language and Cognitive Processes* 25. 149–188.
- Higginbotham, James. 2009. Accomplishments. In James Higginbotham (ed.), *Tense, aspect and indexicality*. 116–125. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199239313.003.0007>.
- Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints on syntactic islands. *Language* 86(2). 366–415. <https://doi.org/10.1353/lan.0.0223>.
- Huang, James Cheng-Teh. 1982. *Logical relations in chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.
- Kaan, Edith, Anthony Harris, Edward Gibson & Phillip Holcomb. 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15(2). 159–201. <https://doi.org/10.1080/016909600386084>.
- Kazanina, Nina & Colin Phillips. 2010. Differential effects of constraints in the processing of Russian cataphora. *Quarterly Journal of Experimental Psychology* 63(2). 371–400. <https://doi.org/10.1080/17470210902974120>.
- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- King, Jonathan W. & Marta Kutas. 1995. Who did what and when? using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7(3). 376–395. <https://doi.org/10.1162/jocn.1995.7.3.376>.
- Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In Peter Culicover & Louise McNally (eds.), *Syntax and semantics 29: The limits of syntax*. 241–279. San Diego, CA: Academic Press.
- Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8. 573–633. <https://doi.org/10.1080/01690969308407588>.
- Kohrt, Annika, Trey Sorensen & Dustin A. Chacón. 2018. The real-time status of semantic exceptions to the adjunct island constraint. *Proceedings of the Chicago Linguistics Society (CLS)* 54.
- Kohrt, Annika, Trey Sorensen, Peter O'Neill & Dustin A. Chacón. 2019. Semantic effects on the processing of filler-gap dependencies into adjuncts and conjuncts. Talk given at the 55th Chicago Linguistics Society.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*. 109–137. Dordrecht: Springer.
- Kutas, Marta & Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (erp). *Annual Review of Psychology* 62. 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Lakoff, George. 1986. Frame semantic control of the coordinate structure constraint. *Proceedings of the Chicago Linguistics Society (CLS)* 22. 152–167.

- Lau, Ellen F., Colin Phillips & David Poeppel. 2008. A cortical network for semantics: (de)constructing the N400. *Nature reviews: Neuroscience* 9. 920–933. <https://doi.org/10.1038/nrn2532>.
- Levin, Beth. 1993. *English verb classes and their alternations: A preliminary analysis*. Chicago: Chicago University Press.
- Maris, Eric & Robert Oostenveld. 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* 164. 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Omaki, Akira, Ellen Lau, Imogen Davidson White, Myles L. Dakan, Aaron Apple & Colin Phillips. 2015. Hyper-active gap filling. *Frontiers in Psychology* 6. 384. <https://doi.org/10.3389/fpsyg.2015.00384>.
- Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82(4). 795–823. <https://doi.org/110.1353/lan.2006.0217>.
- Phillips, Colin. 2013. On the nature of island constraints I: Language processing and reductionist accounts. In Jon Sprouse (ed.), *Experimental syntax and island effects*. 64–108. Cambridge: Cambridge University Press.
- Phillips, Colin, Nina Kazanina & Shani H. Abada. 2005. Erp effects on the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22(3). 407–428. <https://doi.org/10.1016/j.cogbrainres.2004.09.012>.
- Pritchett, Bradley L. 1991. Subjacency in a principle-based parser. In Robert C. Berick, Steven P. Abney & Carol Tenny (eds.), *Principle-based parsing*, 301–345. Dordrecht: Springer.
- R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 2013. Locality. *Lingua* 130. 169–186. <https://doi.org/10.1016/j.lingua.2012.12.002>.
- Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.
- Sassenhagen, Jona & Dejan Draschkow. 2016. Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology* 56. e13335.
- Stowe, Laurie. 1986. Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1(3). 227–245. <https://doi.org/10.1080/01690968608407062>.
- Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Evaraert & Henk van Riemsdijk (eds.), *Blackwell companion to syntax*, vol. 4. 1–51. Hoboken, NJ: John Wiley.
- Traxler, Matthew J. & Martin J. Pickering. 1996. Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35(3). 454–475. <https://doi.org/10.1006/jmla.1996.0025>.
- Truswell, Robert. 2007a. Extraction from adjuncts and the structure of events. *Lingua* 117(8). 1355–1377. <https://doi.org/10.1016/j.lingua.2006.06.003>.
- Truswell, Robert. 2007b. Tense, events, and extraction from adjuncts. *Proceedings of the Chicago Linguistic Society (CLS)* 43. 233–247.
- Truswell, Robert. 2011. *Events, phrases, and questions*. Oxford: Oxford University Press.
- Uriagereka, Juan. 1999. Multiple spell-out. In Norbert Hornstein & Samuel D. Epstein (eds.), *Working minimalism*. 251–282. Cambridge, MA: MIT Press.
- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- de Vincenzi, Marica. 1991. *Syntactic parsing strategies in italian: The minimal chain principle*. Dordrecht: Springer.
- Wagers, Matthew W. & Colin Phillips. 2014. Going the distance: Memory and control processes in active dependency formation. *The Quarterly Journal of Experimental Psychology* 67(7). 1274–1304. <https://doi.org/10.1080/17470218.2013.858363>.
- Yoshida, Masaya, Nina Kazanina, Leticia Pablos & Patrick Sturt. 2014. On the origin of islands. *Language and Cognitive Processes* 29(7). 761–770.