

# The Real-Time Status of Semantic Exceptions to the Adjunct Island Constraint

Annika Kohrt, Trey Sorensen, Dustin A. Chacón  
University of Minnesota

## 1 Introduction

To understand the meaning of a sentence like (1), comprehenders must relate the filler (*the man*) to the gap position in which it is interpreted (marked with an underscore here). Many results have demonstrated that gaps are actively constructed (Crain & Fodor 1985; Stowe 1986; Traxler & Pickering 1996). For instance, Traxler & Pickering (1996) found that reading times increased at the verb *wrote* in the sentence (2a) compared to controls, indicating that comprehenders initially understood *the city* as the object of the verb *wrote*, raising reading times due to the semantic implausibility. Importantly, active gap formation processes appear to be suppressed in syntactic island contexts. For instance, the plausibility mismatch effect observed at the verb *wrote* in (2a) is absent in (2b), because the critical region appears within a syntactic island configuration, i.e., extraction from this position is perceived to be unacceptable, shown in (2c).

- (1) This is the man that Harry said that Dale investigated \_ .
- (2) a. We like the city that the author wrote unceasingly and with great dedication about \_ .  
b. We like the city that the author [ who wrote unceasingly and with great dedication ] saw \_ .  
c. \* We like the book that the author [ who wrote \_ unceasingly and with great dedication ] saw the city.

Traditionally, the unacceptability of sentences like (2c) is posited to follow from syntactic constraints (Ross 1967; Chomsky 1981). The lack of active gap formation in island configurations indicates that these constraints are rapidly and reliably used to guide the comprehender in processing a filler-gap dependency (e.g., Phillips 2006, Yoshida *et al.* 2014). Alternatively, the unacceptability of (2c) may follow from processing limitations (Kluender & Kutas 1993; Kluender 1998; Hofmeister & Sag 2010). Processing a filler-gap dependency requires maintaining a representation of the filler in working memory, and many island configurations are syntactically complex structures that may also independently tax working memory. If so, the comprehender may therefore be unable to deploy active gap formation processes in islands, explaining both the processing behavior and the unacceptability of the sentence in (2c).

We examine these issues through the lens of apparent exceptions to the adjunct island constraint. Many syntactic theories predict that adjunct clauses uniformly are islands (Huang 1982; Uriagereka 1999). However, Truswell (2007,

2011) describes a number of cases where extraction from a non-finite adjunct clause appears to be better than predicted in informal judgments, as shown in (3). For Truswell, extraction is licensed if the filler and the gap are both contained in the syntactic representation of a single event. This means that if the event denoted by the secondary predicate can be identified with the event denoted by the main predicate, and thereby be composed into a “macro-event”, extraction should be licit. For instance, we may take the semantic representation for achievement verbs like *arrive* as something like in (4a) (Higginbotham 1999), in which there is a free event variable denoting the process prior to the moment of arrival. Then, the secondary predicate *whistling* can be identified with the free variable  $e'$ , yielding a representation like in (4b), in which these two events constitute a macro-event. Intuitively, achievement main predicates and activity secondary predicates can more easily compose into a macro-event, licensing extraction.

- (3) a. \* Which tune did Leland dance [ whistling \_ ]?  
 b. Which tune did Leland arrive [ whistling \_ ]?  
 c. Which tune did Leland stand around [ whistling \_ ]?
- (4) a.  $\llbracket \text{arrive} \rrbracket = \lambda x. \lambda e. \text{arrive}(e, x) \wedge \exists E [E = e' \prec e]$   
 b.  $\llbracket \text{arrive whistling} \rrbracket = \lambda x. \lambda e. \text{arrive}(e, x) \wedge \exists E [E = e' \prec e] \wedge \text{whistle}(e')$

In this paper, we examine the status of these apparent exceptions. In Experiment 1, we failed to demonstrate that the choice of main verb influenced extractability from an adjunct clause. In Experiment 2, we investigated whether gaps were actively constructed in adjunct clauses in real-time, and whether this was determined by the semantic characteristics of the main verb. We found that processing time increased when the filler was a plausible argument for the secondary predicate and if the main verb was an achievement. We take this to imply that comprehenders selectively reanalyze a sentence to accommodate an unpredicted gap. Taking the results from Experiments 1 and 2 together, we interpret these results as suggesting that all adjunct clauses are islands, and that the comprehender uses this linguistic constraint to avoid postulating a gap in the adjunct island. However, comprehenders will revise the sentence to a syntactically ill-formed representation to achieve a coherent interpretation. We argue that this willingness to revise to a coherent interpretation drives the improved acceptability of the sentences in (3a), not the differential island status of the adjunct clause. If so, our findings are evidence against resource-based accounts of island phenomena, since we found that comprehenders are capable of constructing a filler-gap dependency into an island configuration, albeit not through predictive (“active”) processes.

## 2 Experiment 1

The goal of Experiment 1 was to determine whether the semantics of the main verb affected the availability of extraction from the adjunct clause. Experiment 1 was an acceptability judgment study, in which participants rated a sentence on a scale from 1 (indicating unacceptability) to 7 (indicating acceptability). The study was conducted on IbexFarm (Drummond 2018). Sixty native English-speaking participants were recruited from Amazon’s Mechanical Turk.

There were 16 target items in our study, and 16 fillers (50% grammatical). We manipulated whether there was a filler-gap dependency ( $\pm$ Wh) and whether the main clause verb permitted extraction ( $\pm$ Extractable). The  $-$ Wh conditions were included as a baseline. The main predicates in the  $+$ Extractable conditions were either achievement verbs or stative verbs, and the  $-$ Extractable conditions were all activities or accomplishments, following the observations in Truswell (2011). The secondary predicates were all activity or accomplishments verbs that were optionally transitive. The materials are exemplified in (5).

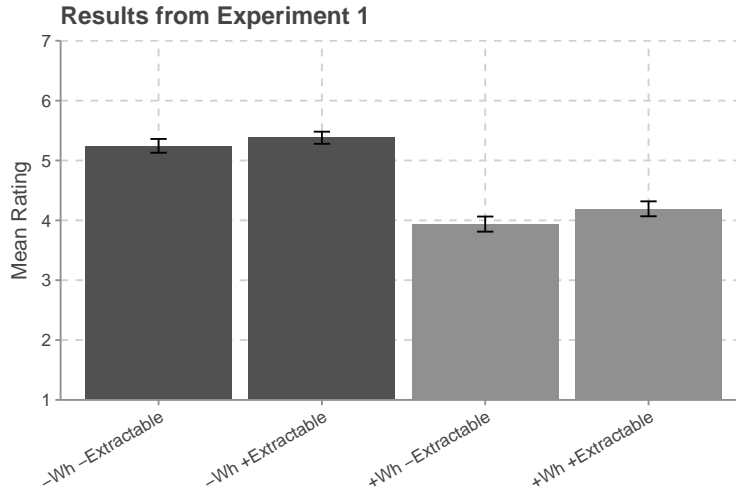
- (5) a. **-Wh,-Extractable**  
 John wondered whether his best friend worked at the office drinking some coffee late this afternoon.
- b. **-Wh,+Extractable**  
 John wondered whether his best friend arrived at the office drinking some coffee late this afternoon.
- c. **+Wh,-Extractable**  
 John wondered which coffee his best friend worked at the office drinking \_ late this afternoon.
- d. **+Wh,+Extractable**  
 John wondered which coffee his best friend arrived at the office drinking \_ late this afternoon.

If extraction from adjunct clauses is allowed when the main verb is an achievement or state, we predict that the ratings for the  $+$ Wh, $+$ Extractable condition should be higher than the  $+$ Wh, $-$ Extractable condition. Conversely, if extraction from adjuncts are uniformly ungrammatical, we expect lower ratings for the two  $-$ Wh conditions across the board, compared to the  $+$ Wh conditions.

The means and standard errors by condition are plotted in Figure 1. For analysis, we submitted the results to mixed-effects model using the lme4 package in R (Bates *et al.* 2016; R Core Team 2017), with both critical factors sum-coded. We included the rating as the dependent variable. For fixed effects, we included  $\pm$ Wh,  $\pm$ Extractable, and their interaction terms. We also included random slopes for  $\pm$ Wh\* $\pm$ Extractable for participants and items<sup>1</sup>. We found a main effect of  $\pm$ Wh ( $\hat{\beta} = 0.62 \pm 0.10, t(27) = 6.7, p < 0.001$ ), but no effect of  $\pm$ Extractable ( $\hat{\beta} = -0.09 \pm 0.09, t(16) = -1.1, p = 0.29$ ) or their interaction term ( $\hat{\beta} = 0.03 \pm 0.08, t(12) = 0.4, p = 0.68$ ). Additionally, pairwise comparisons revealed no difference between the  $-$ Extractable and  $+$ Extractable conditions, either within  $+$ Wh ( $\hat{\beta} = -0.26 \pm 0.23, t(14) = -1.1, p = 0.28$ ) or within  $-$ Wh ( $\hat{\beta} = -0.14 \pm 0.23, t(15) = -0.6, p = 0.56$ ).

To determine whether there was a difference between achievement predicates and stative predicates within the  $+$ Extractable level, we split the  $\pm$ Extractable factor into a new a 3-level factor, VType, with  $-$ Extractable recoded as Activity/Accomplishment, and  $+$ Extractable recoded as State and Achievement, depending on the item. The mean results and standard errors by condition with this new

<sup>1</sup>The structure of our model was  $\text{rating} \sim \text{Wh*Extractable} + (1+ \text{Wh*Extractable}|\text{Item}) + (1+\text{Wh*Extractable}|\text{Participant})$



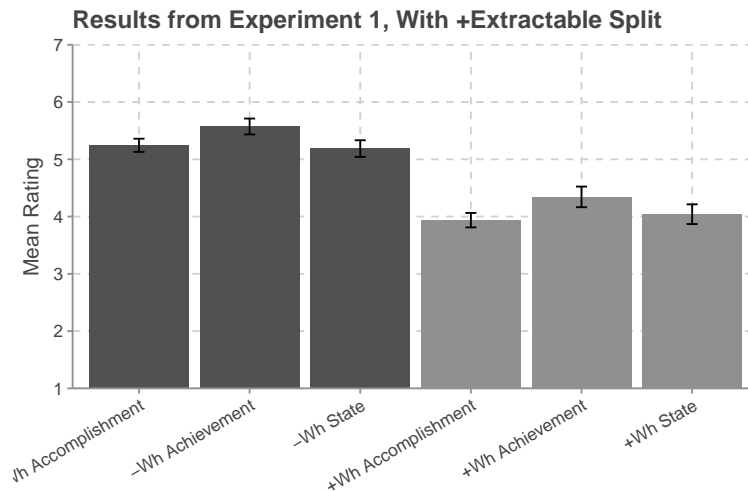
**Figure 1:** Mean ratings by condition from Experiment 1. Error bars represent two standard errors from the mean.

coding are shown in Figure 2. We fit another linear mixed effects model with a similar structure, except substituting the three-level factor VType in for the two-level factor  $\pm$ Extractable.

Again, we find a main effect of  $\pm$ Wh ( $\hat{\beta} = 0.61 \pm 0.10, t(28) = 6.4, p < 0.001$ ). Neither coefficient fit for VType was significant ( $\hat{\beta} = -0.14 \pm 0.12, t(13) = -1.2, p = 0.24$ ;  $\hat{\beta} = 0.14 \pm 0.11, t(12) = 1.3, p = 0.23$ ), nor either of the interaction terms for VType and  $\pm$ Wh ( $\hat{\beta} = 0.04 \pm 0.06, t(632) = 0.67, p = 0.51$ ;  $\hat{\beta} < 0.01 \pm 0.09, t(146) < 0.01, p = 0.97$ ). Finally, although it appears that Achievement verbs in general may be rated more highly than States or Accomplishment verbs, all pairwise comparisons within +Wh ( $\hat{\beta} = -0.24 \pm 0.24, t(23) = -0.99, p = 0.59$ ;  $\hat{\beta} = -0.07 \pm 0.23, t(16) = -0.29, p = 0.96$ ;  $\hat{\beta} = 0.18 \pm 0.26, t(28) = 0.68, p = 0.78$ ) and within -Wh ( $\hat{\beta} = -0.33 \pm 0.24, t(23) = -1.38, p = 0.37$ ;  $\hat{\beta} = -0.23 \pm 0.21, t(20) = -1.07, p = 0.54$ ;  $\hat{\beta} = 0.10 \pm 0.22, t(30) = 0.47, p = 0.88$ ) were insignificant.

Thus, in Experiment 1, we failed to demonstrate that extraction from an adjunct clause was affected by the choice of main predicate. Instead, we found that extraction from adjunct clauses was uniformly penalized. Additionally, although it was neither significant or marginally significant, we did see a small improvement across the board for sentences with main achievement predicates.

As we suggested in the previous section, we propose that the formal judgment results reflect the uniform island status of adjunct clauses. On our proposal, the sentences in (3) diverge in informal acceptability due to the relative ease by which the semantics compose. In Experiment 2, we examine the real-time processing profile of filler-gap dependencies construed into adjunct clauses, and support this second conclusion.



**Figure 2:** Mean ratings by condition from Experiment 1, with +Extractable split into State and Achievement. Error bars represent two standard errors from the mean.

### 3 Experiment 2

The goal of Experiment 2 was to probe whether comprehenders actively constructed gaps in non-finite adjunct clauses, and whether this was modulated by the choice of main verb predicate. Experiment 2 was a moving-window self-paced reading task (Just *et al.* 1982). Participants progressed through a sentence word-by-word by pressing a key, and their reading times per word were measured. After each question, participants were asked to respond to a yes/no comprehension question. Forty-eight native English-speaking members of the University of Minnesota community participated.

We adapted the 16 items from Experiment 1, and included 74 fillers (50% ungrammatical). We used the plausibility mismatch paradigm to probe for active gap formation (Traxler & Pickering 1996). In the plausibility mismatch paradigm, reading times increase if the comprehender actively constructs a filler-gap dependency that is implausible, i.e., there is increased processing difficulty for implausible dependencies compared to plausible dependencies. We manipulated the plausibility of the filler as an argument of the secondary predicate,  $\pm$ Plausible, and the main predicate type,  $\pm$ Extractable. The materials are exemplified in (6).

- (6) a. **+Plausible,-Extractable**  
John wondered which coffee his best friend worked at the office drinking \_ late this afternoon.
- b. **+Plausible,+Extractable**  
John wondered which coffee his best friend arrived at the office drinking \_ late this afternoon.
- c. **-Plausible,-Extractable**  
John wondered which report his best friend worked at the office drinking \_ late this afternoon.
- d. **-Plausible,+Extractable**

John wondered which report his best friend arrived at the office drinking  
\_ late this afternoon.

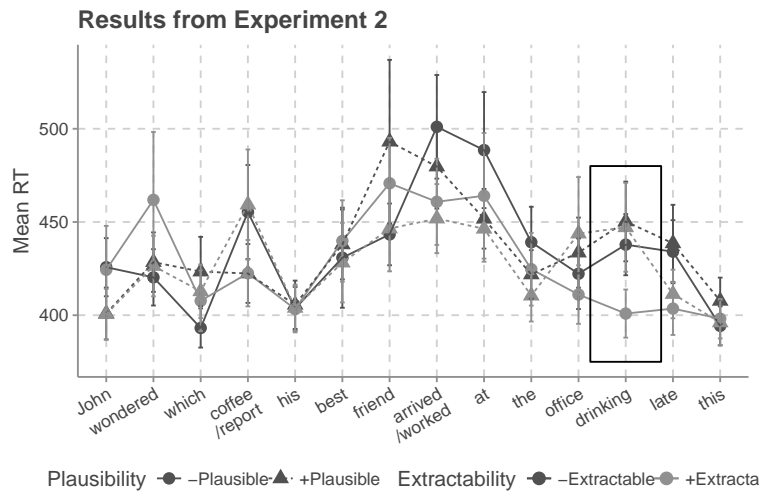
If comprehenders actively expect a gap in adjunct clauses, then we predict increased reading times at the critical region (*drinking*) for the –Plausible conditions. If comprehenders rapidly compute the semantics of the clauses, and then condition their expectation of a gap on the basis of the main clause predicate, then we expect sensitivity to plausibility only within the +Extractable condition. Lastly, as we propose here, if comprehenders selectively reanalyze the sentence to contain a gap in the adjunct clause, but only when the semantics of the sentence is plausible and coherent, then we expect a “reverse” plausibility mismatch effect. That is, we expect increased reading times for the plausible dependencies, reflecting reanalysis, but faster reading times for the implausible dependencies. This prediction leverages the fact that the critical region *drinking* is always an optionally transitive verb, and thus the argument structure of this verb does not force the comprehender to construe the filler as an argument of the verb.

The mean reading times per condition are shown in Figure 3. For analysis, we removed all trials in which the comprehension question was answered incorrectly, which removed 4% of trials. The mean comprehension question accuracy was 90%. Then, we fit a mixed effects model using the lme4 package in R (Bates *et al.* 2016; R Core Team 2017), with the critical manipulations sum-coded. Reading time was included as the dependent variable. We included fixed effects for ±Plausibility, ±Extractability, and their interaction term, and we included random effects for Participant and Item<sup>2</sup>. We found a marginally significant effect Plausibility at the critical region ( $\hat{\beta} = 13 \pm 8, t(621) = 1.7, p = 0.09$ ), but there was no main effect of ±Extractability ( $\hat{\beta} = 9 \pm 8, t(620) = 1.2, p = 0.24$ ), nor any interaction effect ( $\hat{\beta} = -6 \pm 8, t(621) = -0.7, p = 0.47$ ). Planned pairwise comparisons between +Plausible and –Plausible were marginally significant within +Extractable ( $\hat{\beta} = 38 \pm 22, t(621) = 1.7, p = 0.09$ ), but not within –Extractable ( $\hat{\beta} = 15 \pm 22, t(621) = 0.7, p = 0.49$ ).

Next, we split the the ±Extractable factor into the 3-level VType factor, containing Achievement, State, and Accomplishments. The mean reading per condition this split is shown in Figure 4. We fit a model with the same structure as before, except with the VType factor substituted for the ±Extractability factor. Again, we find a marginal main effect of ±Plausibility ( $\hat{\beta} = 15 \pm 8, t(618) = 1.8, p = 0.07$ ), no main effect of either VType coefficient ( $\hat{\beta} = 12 \pm 10, t(617) = 1.2, p = 0.24$ ; ( $\hat{\beta} = -4 \pm 14, t(105) = -0.3, p = 0.78$ )), nor either interaction term with ±Plausibility ( $\hat{\beta} = -7 \pm 10, t(618) = -0.7, p = 0.48$ ;  $\hat{\beta} = 13 \pm 12, t(621) = 1.1, p = 0.27$ ). Planned pairwise comparisons between +Plausible and –Plausible were significant within Achievement ( $\hat{\beta} = 57 \pm 31, t(619) = 1.8, p = 0.07$ ), but not within State ( $\hat{\beta} = 18 \pm 32, t(620) = -0.56, p = 0.58$ ) nor within Accomplishment ( $\hat{\beta} = 15 \pm 22, t(618) = -0.7, p = 0.49$ ). Thus, in Experiment 2, we found that comprehenders were sensitive to the plausibility of a filler as a possible

---

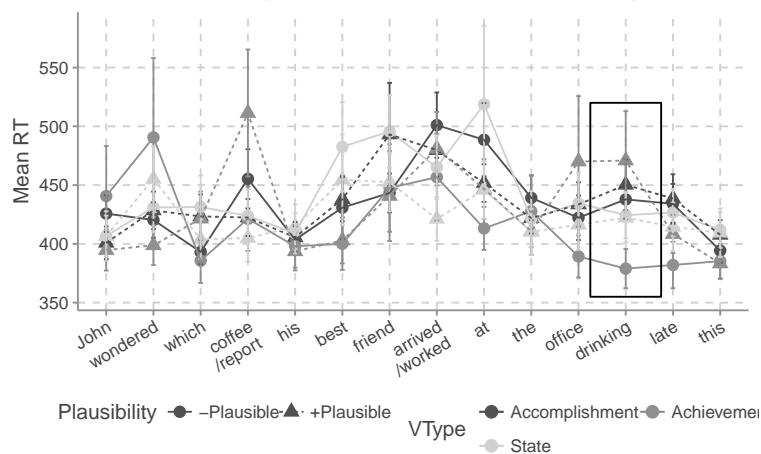
<sup>2</sup>The structure of this model was reading time  $\sim$  Plausibility\*Extractability + (1|Item) + (1|Participant)



**Figure 3:** Mean reading times by region and condition from Experiment 2. Error bars represent two standard errors from the mean.

argument for the secondary predicate, but only when the main verb was an achievement. Furthermore, this was reflected as a “reverse plausibility mismatch effect”, i.e., plausibility resulted in increased processing difficulty.

**Results from Experiment 2, With +Extractable Split**



**Figure 4:** Mean reading times by region and condition from Experiment 2, with +Extractable split into State and Achievement. Error bars represent two standard errors from the mean.

#### 4 Conclusion

The goal of this study was to determine whether the acceptability of extraction from adjunct clauses was determined by the semantics of the main predicate, and whether comprehenders’ expectation for a gap was affected by the semantics of the main

predicate. We found a somewhat perplexing set of results. First, in Experiment 1, we found that extraction from adjunct clauses was generally lower, regardless of the choice of main clause predicate. However, we found a small, statistically insignificant improvement for Achievement verbs in our post-hoc analysis, but this was not specific to sentences with extraction. Then, in Experiment 2, we found that reading times were fastest in conditions where the main clause predicate was an achievement, and the unresolved filler was not a suitable argument for the secondary predicate. All other reading times were increased compared to this condition, although the results were marginally significant.

To account for our results, we maintain the conservative generalization that adjunct clauses are islands, regardless of the semantics of the main clause predicate. This is reflected in offline judgments, as in Experiment 1. In real-time processing, this is reflected as a reluctance to actively postulate a gap in the adjunct clause, yielding no (traditional) plausibility mismatch effect. However, we propose that comprehenders revise the sentence shortly after encountering the secondary predicate, based on the semantics of the predicates. Following Truswell (2007, 2011), we propose that achievement verbs have a free event variable, as in (4a). The comprehender can easily identify this event variable with the secondary predicate, meaning that ratings for secondary predication with achievement verbs are improved and processing is quicker. However, we propose that there is no free event variable in the semantics of a stative or accomplishment predicate. Thus, composing with a secondary predicate requires coercion, which independently increases processing difficulty (Pylkkänen & McElree 2006). This results in a slight preference for secondary predication with achievement verbs, and faster reading times. The representations we have in mind for the three predicate types are sketched in (7).

- (7) a.  $\llbracket \text{arrive} \rrbracket = \lambda x. \lambda e. \text{arrive}(e, x) \wedge \exists E [E = e' \prec e]$   
 b.  $\llbracket \text{work} \rrbracket = \lambda e. \text{work}(e)$   
 c.  $\llbracket \text{stand around} \rrbracket = \lambda x. \lambda e. \text{standaround}(e, x)$

Secondly, to explain the reverse plausibility mismatch effect that we found in Experiment 2, we suggest that the comprehender chooses to reanalyze to an ill-formed sentence to maintain a coherent semantic interpretation. If the main verb is an achievement, e.g., *arrived*, then semantically integrating *drinking* is facilitated. However, if the filler was a semantically suitable argument for the secondary predicate, e.g., *coffee*, then the comprehender “disobeys” the adjunct island constraint to maintain a coherent interpretation, i.e., one in which coffee is drunk. This results in longer reading times for the sentence that is reported to be acceptable in informal judgments, due to reanalysis. Since there is no well-formed gap later in the sentence, reanalyzing to a coherent but ill-formed interpretation may be preferred to abandoning the filler-gap dependency entirely. Importantly, this is gated by the semantics of both the filler-gap dependency and the two predicates, i.e., this is conditioned on global semantic coherence.

Conversely, if the filler is an unsuitable argument for the secondary predicate, but the main verb is an achievement, reading times were faster. We argue that this reflects the ease of composition between the two predicates, and failure to reanalyze. That is, comprehenders were not tempted into reanalyzing the sentence to



permit the implausible argument to be interpreted as the object of the secondary predicate (reports are rarely drank). Additionally, for the accomplishment and state verbs, there was no sensitivity to the plausibility of the filler.

To further support this analysis, we are conducting further behavioral tasks to probe for the role of event structure in the construction of filler-gap dependencies, and we are using ERP measures to distinguish the two processes that we argue underlie apparent extraction from an adjunct clause (relating main and secondary predicates and revising in accordance to maintain coherence).

These results have significance for the debate over the nature of island effects. Our results suggest that comprehenders generally have sufficient working memory resources to construct filler-gap dependencies into islands. Thus, the unacceptability of extraction from adjunct clauses (as demonstrated in Experiment 1) cannot be due to limitations on working memory capacity. Instead, we propose that grammatical constraints block this interpretation (e.g., Phillips 2006; Yoshida *et al.* 2014). However, our analysis suggests a more nuanced view. For us, predictive processes are constrained by linguistic constraints, but not revision processes. This is why comprehenders are capable of building an ill-formed filler-gap dependency “bottom up” through reanalysis, in order to maintain a coherent semantic/conceptual representation. This finding has analogs in the literature on resumptive pronouns. Resumptive pronouns are rated to be unacceptable in formal judgments, but are perceived to be acceptable in informal judgments. Similarly, they appear to facilitate “bottom-up” completion of an otherwise unacceptable dependency (Heestand *et al.* 2011; Hofmeister & Norcliffe 2013; Chacón 2015). Lastly, our proposal relies on the comprehender noticing a strong lexical association between the plausible filler (e.g., *coffee*) and the structurally inaccessible adjunct (e.g., *drinking*) to guide this reanalysis process. Thus, there are parallels with “Good Enough” models of sentence processing (Christianson *et al.* 2001; Ferreira & Patson 2007) and “Self-Organized Parser” models (Tabor *et al.* 2004; Kukona *et al.* 2014), which both allow semantic and lexical information to influence syntactic processing to some degree.

Similarly, our results may be informative for the growing literature on the nature of acceptability judgments (Sprouse 2008; Bader & Hussler 2010; Ackerman *et al.* 2017). On our proposal, formal judgments and predictive processes more closely reflect grammatical constraints, whereas informal judgments and revision processes are susceptible to building ill-formed but coherent representations. However, there is still an unexplained mismatch between informal and formal judgments. The results from Experiments 1 and 2 specifically pinpoint achievement predicates as facilitating extraction from adjunct clauses, whereas informal judgments indicate that achievement and stative predicates both allow extraction (as in 3b and 3c). We did not explicitly set out to compare these different classes of predicates, but we hope to more carefully compare these factors in future studies.

## References

- Ackerman, L., M. Frazier, & M. Yoshida. 2017. Resumptive pronouns can ameliorate illicit island extractions. *Linguistic Inquiry* To appear.
- Bader, M., & J. Hussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics* 46.273–330.
- Bates, D., Maechler Martin, & S. Walker. 2016. Package ‘lme4’.
- Chacón, D. A. *Comparative Psychosyntax*. University of Maryland dissertation.
- Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
- Christianson, K., A. Hollingworth, J. F. Halliwell, & F. Ferreira. 2001. Thematic roles assigned along the garden path linger. *Cognitive Psychology* 42.368–407.
- Crain, S., & J. D. Fodor. 1985. How can grammars help parsers? In *Natural language parsing*, ed. by D. R. Dowty, L. Karttunen, & A. M. Zwicky, 94–128. Cambridge: Cambridge University Press.
- Drummond, A. 2018. IbeXFarm. <http://spellout.net/ibexfarm>. Retrieved 17 Apr 2018.
- Ferreira, F., & N. D. Patson. 2007. The ‘Good Enough’ Approach to Language Comprehension. *Language and Linguistics Compass* 1.71–83.
- Heestand, D., M. Xiang, & M. Polinsky. 2011. Resumption Still Does Not Rescue Islands. *Linguistic Inquiry* 42.138–152.
- Higginbotham, J. 1999. Accomplishments. In *Proceedings of the Nanzan GLOW: The Second GLOW Meeting in Asia*, ed. by M. Saito, 131–139, Nagoya. Nanzan University.
- Hofmeister, P., & E. Norcliffe. 2013. Does resumption facilitate sentence comprehension? In *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, ed. by P. Hofmeister & E. Norcliffe, 225–246. Stanford: CSLI Publications.
- Hofmeister, P., & I. A. Sag. 2010. Cognitive constraints and island effects. *Language* 86.366–415.
- Huang, C.-T. J. *Logical relations in Chinese and the theory of grammar*. MIT dissertation.
- Just, M. A., P. A. Carpenter, & J. D. Woolley. 1982. Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General* 111.228–238.
- Kluender, R. 1998. *On the distinction between strong and weak islands: A processing perspective*. San Diego, CA: Academic Press.
- Kluender, R., & M. Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8.573–633.
- Kukona, A., P. W. Cho, J. S. Magnuson, & W. Tabor. 2014. Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. *Journal of Experimental Psychology: Learning Memory and Cognition* 40.326–347.
- Phillips, C. 2006. The Real-Time Status of Island Phenomena. *Language* 82.795–823.
- Pylkkänen, L., & B. McElree. 2006. The syntax-semantic interface : On-line composition of sentence meaning. In *Handbook of Psycholinguistics (Second Edition)*, ed. by M. J. Traxler & M. A. Gernsbacher, 539–579. New York: Elsevier.
- R Core Team. 2017. R: A Language and Environment for Statistical Computing.
- Ross, J. R. *Constraints on variables in syntax*. MIT dissertation.
- Sprouse, J. 2008. The Differential Sensitivity of Acceptability Judgments to Processing Effects. *Linguistic Inquiry* 39.686–694.
- Stowe, L. 1986. Parsing WH-Constructions: evidence for on-line gap location. *Language and Cognitive Processes* 1.227–245.
- Tabor, W., B. Galantucci, & D. Richardson. 2004. Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language* 50.355–370.
- Traxler, M. J., & M. J. Pickering. 1996. Plausibility and the Processing of Unbounded Dependencies: An Eye-Tracking Study. *Journal of Memory and Language* 35.454–475.
- Truswell, R. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117.1355–1377.
- Truswell, R. 2011. *Events, Phrases, and Questions*. Oxford: Oxford University Press.
- Uriagereka, J. 1999. Multiple Spell-out. In *Working Minimalism*, ed. by N. Hornstein & S. D. Epstein, 251–282. Cambridge: MIT Press.
- Yoshida, M., N. Kazanina, L. Pablos, & P. Sturt. 2014. On the origin of islands. *Language, Cognition and Neuroscience* 29.761–770.