Alternatives on demand and locality: Resolving discourse-linked *wh*-phrases in sluiced structures.

J. A. Harris: Department of Linguistics, University of California, Los Angeles, USA

jharris@humnet.ucla.edu (310) 825-0634

Abstract

Previous studies have observed a tendency to associate the remnant (e.g., *who*) of ambiguous sluicing ellipsis with the closest / most local correlate (*someone*) in the matrix clause, as in *Somebody said Fred fired someone, but I don't know who.* I present the results of three experiments investigating the interplay between locality and the discourse status of potential correlates. The studies exploit the discourse-linking property of *which*-phrases in ambiguous sluiced sentences, like *A teacher scolded Max or Dotty, but I can't remember which one*, to explore whether the preference for more local correlates is modulated by the discourse status of the potential correlates. I propose a discourse economy constraint (*Alternatives on Demand: Avoid positing new discourse alternatives without evidence*), which interacts with structural constraints like locality. Evidence from several questionnaire studies, as well as three online self-paced reading studies, supports the predictions of a sentence processing model in which the discourse status of items in memory immediately impacts the retrieval of a correlate for the remnant of sluicing ellipsis and related constructions. In addition, the time point at which the interaction between processing biases appears is shown to depend on the strength or diagnosticity of the retrieval cues in *which*-phrase.

1 Introduction

During online comprehension, the sentence processor must negotiate a great many disparate types of information to produce a sensible output, plausibly engaging parsing biases or heuristics to produce an initial grammatical structure. Some of these biases are perhaps driven by structural economy factors, in which syntactically less complex sentence parses are preferred over others (Frazier, 1987). Other processing biases may involve discourse factors including accessibility, familiarity, and salience (Arnold 2010, for review). The interplay between such factors should be especially evident in cases of form-meaning mismatch, such as ellipsis, in which the meaning of an expression must be recovered by retrieving or inferring prior linguistic material.

In three experiments, I examine how meaning is recovered in clausal ellipsis and related ellipsis cleft structures, showing how two biases interact in the interpretation process. The studies pit a discourse-economy principle militating against positing potentially unnecessary discourse referents, against a well-established structural bias associated with the interpretation of clausal ellipsis. I concentrate on constructions known as 'sluices,' defined as clausal ellipsis of a constituent question (Ross, 1967; Chung et al., 1995; Romero, 1998; Merchant, 2001; van Craenenbroeck, 2010, among others). In (1a), for example, the verb *left* has been elided so that only the *wh*-element (in this case *who*) remains as the remnant of the ellipsis. I assume that the processor must recover the 'missing' constituent Δ in order to completely interpret the sentence.

(1) a. Somebody left – guess [$_{CP}$ who$_1$ Δ]

b. Somebody left – guess [$_{CP}$ who$_1$ [$_{IP}$ t_1 ~~left~~]]

Sluicing typically associates the remnant (*who*) with a correlate (*somebody*) in the antecedent clause (*Somebody left*). Although there are many kinds of possible correlate-remnant pairs, the correlate evokes alternatives of the same semantic type as the remnant (e.g., Barros & Vicente, 2016). Usually, the correlate is an indefinite like *somebody* (1) or *a student* (2a), or a weak definite (like *the doctor*).[1] A disjunction (*John or Bill*) may also serve as a correlate to the remnant of a sluice (Chung et al., 1995), which is expected if both indefinite nouns and disjunctions evoke alternatives (e.g., Kratzer & Shimoyama, 2002; Alonso-Ovalle, 2006) or introduce issues to be resolved by the discourse (AnderBois, 2014). In addition, some kinds of *wh*-element place constraints on their correlates: the remnant *which (one)* is licensed for indefinites and disjunctions, but not for other indefinites (*someone*, as in (3a)). Conversely, a *who* remnant is relatively unnatural with indefinite or disjunctive correlates, unless followed by *else* (3b); see Dayal & Schwarzschild (2010) for discussion.

(2) a. A student laughed, but I can't say which (one / student).

b. John or Bill laughed, but I can't say which (one / of them).

(3) a. * Someone laughed, but I can't say which (one).

b. A student / John or Bill laughed, but I can't say who *(else).

Which-phrases impose further requirements on their antecedents, and are said to be 'discourse-linked' to salient discourse entities (Pesetsky, 1987). In addition, the content contained within the inner restrictor of the *wh*-element, such as the pronoun *one* in *which one*, or a nominal *student* in *which student*, intuitively contributes to the interpretation of the sluice. The present paper

[1] Chung et al. (1995) also note that the remnant can correspond to an adverbial like *where*, *why*, or *how*, and that a variable *x* related to the correlate can "sprout" within the ellipsis site (*He ate, but didn't tell me what <he ate x>*). Such cases will not be discussed here.

explores the relationship between a structural bias governing the interpretation of sluiced sentences, and the discourse biases imposed by the *which*-phrase in the remnant. I present evidence that these biases interact during online processing, and propose that the results support a model in which both structural and discourse information contribute to the resolution of the remnant in real time.

1.1 Processing sluicing and the Locality Bias

Sluicing and related ellipsis constructions have received a surge of attention in recent psycholinguistics literature, concentrating on the effects of parallelism (Frazier & Clifton, 1998; Carlson 2002; Dickey & Bunger, 2011), focus sensitivity (Frazier et al., 2006; Carlson et al., 2009), and the mechanisms of retrieval (Frazier & Clifton, 2005; Martin, 2010; Martin & McElree, 2008, 2011; Poirier et al., 2010; Harris, 2015). Following Harris & Carlson (2016), I assume that the processor must complete three basic tasks when interpreting clausal ellipsis (4).

(4) **Basic tasks of the processor in clausal ellipsis:**

1. Parse the remnant by constructing the appropriate phrase structure for the remnant given the input.
2. Locate the correlate, if any, from the antecedent clause and pair it with the remnant.
3. Construct the elided phrase by regenerating or copying a structure at Logical Form.

I am primarily concerned with the second task in (4) in which a suitable correlate for the remnant must be identified. Although the cases of sluicing discussed so far have all been unambiguous in terms of pairing the correlate with the remnant, antecedents to sluiced sentences may have

multiple semantically appropriate constituents that could serve as potential correlates (6a). In such cases, the remnant preferentially associates with the most local constituent of the appropriate type (Frazier & Clifton, 1998; Carlson et al., 2009), a preference sometimes referred to as the Locality Bias (Harris, 2015; Harris & Carlson, 2016, in press).

(5) **Locality Bias.** Contrast the remnant with the nearest constituent (of the appropriate type) in the preceding clause.

Several studies confirm the central predictions of the Locality Bias. In a self-paced reading paradigm, Frazier & Clifton (1998) found that unambiguous sluices with subject position correlates (6b) were read slower than ambiguous counterparts with object correlates (6a).

(6) a. Somebody claimed that the president fired someone, but nobody knows who.
b. Somebody claimed that the president fired Fred, but nobody knows who.

In an auditory questionnaire, Carlson et al. (2009) found additional support for the Locality Bias by varying focus placement in sentences like *Alice insulted Bill, but I don't know who else*. When only the subject noun (*Alice*) was pitch accented, subjects tended to choose the subject as the correlate (40% object correlates) for the remnant (*who else*); otherwise, it was strongly biased towards the object noun (*Bill*), including cases in which both the subject and object were accented (~ 80% object correlates).

In an eye tracking study, Harris (2015) found that local correlates (*some wines*) to remnants in unambiguous sluicing ellipsis facilitated reading time over non local correlates

(*some tourists*), as in (7). In addition, the re-reading penalty for non local correlates was greater when the non-correlate distractor noun was in the preferred local position and was similar in grammatical number (*the tourists* or *the wines*) with the correlate target.

(7) a. The tourist(s) sampled *some wines*, but I've forgotten which ones / wines … (Local)

b. *Some tourists* sampled the wine(s), but I've forgotten which ones / tourists … (Non Local)

The results were interpreted in terms of similarity-based interference effects, predicted by cue-based models of retrieval (e.g., Lewis, Vasishth, & Van Dyke, 2006). Assuming that the pairing between the correlate and the remnant is similar to general retrieval processes, distractor items that share features with the target should slow retrieval as the retrieval cue is "overloaded" (Watkins & Watkins, 1975; Van Dyke & Johns, 2012). Harris (2015) described the effects in terms of variable cue "diagnosticity", in that remnants that specify or distinguish the correlate from distractors are subject to reduced interference effects, facilitating the speed and accuracy of correlate-remnant pairings. In example (7), Nominal restrictors (*which wines*, *which tourists*) and pronouns that agree with only one noun (*which ones*) provide strongly diagnostic cues for the target remnant, whereas remnants with pronouns (*which one*) that are compatible with either noun in the antecedent clause are weakly diagnostic for retrieval. Remnants with strongly diagnostic cues showed a reduced Locality effect compared to those with weakly diagnostic cues, indicating that correlate location modulates interference during retrieval.

How can we explain the Locality Bias? One possibility is that the processor simply retrieves the most recently activated element to minimize demands on working memory. This initially appealing possibility, however, is not supported by recent research. In a cross-modal

priming study, Poirier et al. (2010) presented printed targets related to the subject (*the handyman*) or the dative object (*the programmer*) distractors at two probe points in auditory sentences like (8). The first probe location appeared immediately after the offset of the remnant $*_1$; the second probe point was located 500ms downstream $*_2$.

(8) The handyman threw a book to the programmer but I don't know which book $*_1$ and no one $*_2$ else seems to know.

Although there was an advantage for targets related to the object at the second probe point, there was no difference between subject and object related targets at the remnant, suggesting that both antecedents were equally accessible at the remnant. Another possibility, due to Carlson et al. (2009), is that default linguistic focus placement produces the Locality Bias, an idea that coheres well with theoretical literature positing focus as an essential ingredient in the interpretation of ellipsis (Romero 1998; Merchant 2001). On this account, it is assumed that the antecedent clause bears a main sentence accent on its most deeply embedded constituent by default (Selkirk, 1984; Cinque, 1993). As the object is the most embedded constituent in canonical SVO clauses, default accent would facilitate access of the object when locating a correlate to the remnant of ellipsis (though see Harris & Carlson, in press, for additional discussion).

1.2 Discourse-linking and alternatives on demand

As observed by Pesetsky (1987), certain *wh*-phrases like *which* impose specific constraints on discourse accessibility. Using the phrase *which N*, as in *which ones* in (9a), presumes that the set of felicitous answers is constrained to a salient set of men provided by the

discourse. This is known as d(iscourse)-linking, as the *which* phrase is said to be "linked" to the entities in the discourse. The d-linking property appears to be unique to, or at least particularly strong with, *which* phrases, as the more general *who* element in (9b) may be uttered even if the speaker has no specific set of men previously mentioned in mind.

(9) a. Some men entered the room. Which (ones) did Mary talk to?

b. Some men entered the room. Who did Mary talk to?

The d-linking requirement of *wh*-phrases appears to manifest immediately. Frazier & Clifton (2002) hypothesized that "the d-linked phrase requires the postulation of a discourse entity, while *who* does not. Because pronouns seem to prefer antecedents in a discourse representation, this makes *which*-N relatively more available" (see also Frazier et al. 1996). In support of this hypothesis, they found that d-linked *wh*-phrases were more likely to be taken as antecedents to pronouns than were non-d-linked interrogatives. In a reading time follow up, regions containing the pronoun were read faster when following a d-linked *wh*-phrase. These results are expected if the sentence processor immediately accesses a model of the discourse when processing both pronouns (Cloitre & Bever, 1988; Garnham et al., 1995) and d-linked *wh*-phrases, creating discourse entities for *which* phrases when no suitable ones exist in the discourse.

Although listeners are sensitive to new and given information status in discourse in general (e.g., Kaiser and Trueswell, 2004; Wolter et al., 2011), the preference for alternatives given in previous discourse may be particularly strong for d-linked *wh*-phrases. The pattern follows from a general constraint, which I coin 'Alternatives on Demand' – which, simply put,

militates against positing potentially unnecessary discourse alternatives into the discourse representation.

(10) **Alternatives on Demand (AD).** Posit discourse alternatives only when required, e.g., by focus placement.

The AD principle embodies a simple precept: keep the discourse representation unencumbered by potentially irrelevant contrast sets (see also Sedivy, 2002; Beaver & Clark, 2008; Carlson, 2013, and especially Kim et al., 2015). This assumption that generating alternatives is computationally demanding is standard within current theories of focus semantics (such as alternative semantics; Rooth, 1985) in which a covert operator ~ induces a set of alternative propositions that is formed by replacing the focus-marked element $[.]_F$ with each alternative individually. For example, if our alternative set ALT contained the individuals Alex, Ben, and Claude, applying the operator ~ to *Abbey loves* $[Ben]_F$ would result in a set of propositions: {Abbey loves Alex, Abbey loves Ben, Abbey loves Claude}. Generating focus alternatives therefore involves three major steps: (i) positing a covert sentential operator ~ in the correct position, (ii) determining the set of alternatives ALT from context, and (iii) substituting elements of ALT for the focus-marked element to derive the set of alternative propositions. Thus, interpreting a sentence with discourse alternatives plausibly involves multiple sources of computational complexity, the most unconstrained of which is step (ii) in which the ALT set is determined from context. The AD principle can be interpreted as the injunction to reduce computational complexity by selecting a set of alternatives ALT that is already given by previous discourse or directly from the sentence itself.

The following studies concentrate on the interpretation of sluice structures with explicit and implicit alternatives. By explicit alternatives, I mean alternatives explicitly provided by the sentence itself, so that they are discourse given or discourse old. The studies below use disjunctions, which present the relevant discourse alternatives directly in the sentence. In *John talked to Bob or Sue*, for example, the disjunction *or* introduces two, non-mutually exclusive, possibilities*: that John talked to Bob* or *that John talked to Sue*. By implicit alternatives, I mean alternatives that must be generated by semantic or discourse operators. I utilize the semantics of indefinite descriptions, like *a guest,* for implicit alternatives, which may either be interpreted as a singular entity, or a set of individuals satisfying the property of being a guest. The basic prediction of Alternatives on Demand is that computing potentially unmotivated alternatives should be avoided if possible. This preference should be especially strong with d-linked *which* phrases, which independently prefer given antecedents. Therefore, a *which* remnant should prefer a disjunction to an indefinite as its correlate, following a general discourse preference for explicit over implicit alternatives.

2 Experiment 1

Two offline interpretation questionnaire studies in the following section address the interaction between the Locality Bias and Alternatives on Demand in pairing a correlate with the remnant, i.e., step 2 in the procedure sketched in (4). In general, disjunctions should be preferred to indefinites as correlates in sluicing, but the strength of the preference may be affected by how local the potential correlate is. The studies also address the extent to which an ambiguous singular pronoun (*one*) in the restrictor of the remnant influences interpretation compared to a bare *which* remnant without an overt restrictor.

2.1 Experiment 1A: Interpretation questionnaire

Thirty-six native speakers of English were recruited on Amazon's Mechanical Turk, an Internet-based service where workers complete small tasks, and were compensated with $5. Only individuals who had performed at least 50 previous questions and received a 98% approval rating or above were permitted to participate in the experiment. Locality, the position of the disjunction (*Local* vs. *Non Local*), was crossed with whether the d-linked remnant (*which*) contained the pronoun *one* (Pronoun) or not (Null), as shown in (11). After reading the sentences, participants chose one of two answers to an interpretation question (12).

(11) a. A guest talked to Bill or Sue, but I don't remember which. (Local - Null)

b. A guest talked to Bill or Sue, but I don't remember which one. (Local - Pronoun)

c. Bill or Sue talked to a guest, but I don't remember which. (Non-Local - Null)

d. Bill or Sue talked to a guest, but I don't remember which one. (Non-Local - Pronoun)

(12) *Interpretation question:* What don't I remember?

i. I don't remember which of Bill or Sue it was. (Disjunction-antecedent)

ii. I don't remember which of the guests it was.

Disjunction-antecedent responses will be treated as the response type of interest in the statistical analyses of forced-choice questions in all experiments that follow. Presentation order of items was randomized, as was the presentation of forced choice answers. Means and standard errors are provided in Table 1. Data were analyzed as a linear mixed effects logistic regression model (Baayen et al., 2008; Jaeger, 2008), using the lme4 package (Bates & Maechler, 2009)

with sum (deviation) contrast coding with maximal random effects structures: by-subjects and by-items random slopes and intercepts (following Barr et al., 2013). Table 2 presents the statistical models for all experiments reported in Experiment 1.

Experiment	Remnant type	Local disjunction	Non Local disjunction	Difference: Locality bias
1A	Null	91% (2)	67% (4)	24%
	Pronoun	88% (3)	72% (4)	16%
1B	Null	81% (3)	68% (4)	13%
	Pronoun	76% (3)	61% (4)	15%

Table 1. Experiment 1: Percentages of Disjunction-antecedent choices (Experiment 1A) and disjunction completions (Experiment 1B). Standard errors in parentheses.

Experiment	Parameters	Estimate	Std. Error	Wald *Z*	*p* estimate
1A	(Intercept)	3.083	0.487	6.336	< 0.001
	Non Local	-1.535	0.423	-3.629	< 0.001
	Pronoun	-0.852	0.222	-3.834	< 0.001
	Non Local x Pronoun	0.436	0.227	1.923	0.054
1B	(Intercept)	5.370	1.422	3.776	< 0.001
	Non Local	-2.537	0.669	-3.79	< 0.001
	Pronoun	0.091	0.433	0.210	0.834
	Non Local x Pronoun	-0.149	0.364	-0.408	0.683

Table 2. Experiment 1: Fixed effects from logistic linear mixed-effects regression models for both questionnaire studies in Experiment 1.

Analysis revealed a general preference for disjunction correlates (M = 78%), as well as a main effect of disjunction position: there were significantly more Disjunction-antecedent responses in the Local (object disjunction) condition (M = 90%, SE = 2) than the Non-Local (subject disjunction) condition (M = 65%, SE = 3), z = -3.07, $p < 0.001$. In addition, a main effect of Remnant type was observed, such that remnants with the pronoun *one* received fewer Disjunction-antecedent responses (M = 75%, SE = 3) than those without (M = 79%, SE = 2), z = 3.83, $p < 0.001$. There was also a marginal interaction: the effect of pronoun was marginally greater for Local disjunctions (*diff* = 5%) than Non Local disjunctions (*diff* = 3%), z = 1.92, p =

0.05, which may indicate model over-fitting given the modest difference between conditions. The robustness of the effect is explored in a second questionnaire study within a different paradigm.

2.2 Experiment 1B: Fill in the blank

Instead of choosing between two fixed alternative interpretations, participants completed the sentence with a single word. This method was adopted in order to avoid any influence from the response paraphrase in (12) and to determine whether the previous effects would persist in a more open-ended task. A distinct thirty-six subjects were recruited on Amazon's Mechanical Turk as in the previous experiment. Materials were identical to the first 20 sentences from the previous experiment, except that the disjunction "or" was replaced with a blank. Participants were instructed to fill in the blank with the single word that best completed the sentence. The manipulation crossed position of the blank (*Local* vs. *Non Local*) with Remnant type: the presence or absence of the pronominal *one* in the *which* phrase. All participants provided either a disjunction *or*, or a conjunction *and,* in the blank.

(13) a. A guest talked to Bill ____ Sue, but I don't remember which (one). (Local blank)

b. Bill ____ Sue talked to a guest, but I don't remember which (one). (Non Local blank)

Participants generally favored disjunctions (72% disjunction preference across conditions), and were also more likely to provide a disjunction when the blank appeared in Local (M = 78%, SE = 3) compared to Non Local position (M = 65%, SE = 2), z = -3.79, $p < 0.001$. The presence of a pronoun did not affect the response type, nor was there an interaction ($z < 1$). Although the

disjunction does not itself disambiguate the interpretation of the sluice, the results are highly compatible with the view that subjects prefer disjunctions in the local object position, because it allows them to avoid postulating alternatives for the non local subject noun (*a guest*).

2.3 Discussion

Two questionnaire studies tested the relationship between two constraints: Alternatives on Demand, which favors interpretations that avoid computation of alternative semantic values, and the Locality Bias, a previously-observed tendency to resolve ambiguous ellipsis to the most local correlate of the appropriate type of the preceding clause. By Alternatives on Demand, disjunctions should in general be preferred as correlates to d-linked remnants of sluicing ellipsis over indefinite counterparts, since disjunctions provide overt alternatives – i.e., the disjuncts themselves. By Locality, the processor should favor resolutions to the object, though whether it does so for reasons of focus or accessibility is still unclear. The central prediction was confirmed across both paradigms: resolution of a d-linked remnant *which (one)* is affected by the position of the preferred antecedent. When the two constraints converge, the processor favors the disjunctive object noun (at a rate of 75–90%). When they do not, an indefinite object noun becomes much more tempting (at a rate of 55–65%), as indicated from the Non local conditions above. The resolution biases are remarkably similar to those observed in Carlson et al. (2009), who manipulated accent placement in ambiguous sluices. The two studies may be related in that focus, as signaled by pitch accent placement, generates alternatives. The main difference can be described in terms of how ALT, the set from which alternatives are taken, is determined. Whereas ALT is *contextually* determined in the case of focus, it is *overtly* determined in the case of disjunction.

Although it is not yet clear how the presence of pronominal *one* affects these preferences – given that the effect of the pronoun observed in Experiment 1A was numerically small and failed to reach significance in Experiment 1B, the direction of the results indicates that *one* could not explain the preference for explicit alternatives. In all, the central results clearly support a processing model in which structural and discourse information are taken into consideration during the retrieval of correlates in sluicing ellipsis. Further, the processor does not simply default to just one of the structural or discourse biases when in conflict.

The following two experiments probe whether the conflicting preferences produce costs in online processing. On the one hand, it might be that one or both preferences constitute entirely offline, post-processing strategies, and only operate to resolve cases of ambiguity. On the other, Locality and Alternatives on Demand might guide online processing decisions, and induce processing difficulties when both cannot be met.

3 Experiment 2

3.1 Experiment 2A

The experiment was designed to investigate whether Locality and Alternatives on Demand operate during online sentence processing, and the extent to which the diagnosticity of retrieval cues in the remnant affect the retrieval of an appropriate correlate (Harris, 2015). Forty-eight native speakers of English from the Claremont Colleges participated in the study for $10. Materials consisted of 30 sextets in a 3 × 2 design crossing Restrictor and Locality. The Restrictor factor describes the restrictor content in the remnant along three levels: *Singular*, in which the remnant contained the pronoun *one*, *Plural*, in which the remnant contained *of them*, and *Nominal*, in which the remnant contained the NP from the indefinite, e.g., *guest* in (14)

below. The second factor of Locality comprised two levels: *Local* and *Non Local*. Locality affects Nominal conditions differently than other conditions: for the Nominal condition, the Locality Bias is satisfied by an object indefinite – the Non Local disjunction of (14b), whereas for Singular and Plural conditions, the Locality Bias is satisfied by a Local disjunction (14a). Sentences were presented in a cumulative self-paced moving window format. The region of central interest contained the disambiguating restrictor, Region 6. A forced-choice question was presented after each trial (15), similar to Experiment 1A. Linger (http://tedlab.mit.edu/~dr/Linger/) was used to present materials and record responses.

(14) a. $|_1$ A guest $|_2$ talked to $|_3$ Bill or Sue, $|_4$ but I don't $|_5$ remember which $|_6$ one / of them / guest, $|_7$ because it was $|_8$ so long ago.| (Local disjunction)

b. $|_1$ Bill or Sue $|_2$ talked to $|_3$ a guest, $|_4$ but I don't $|_5$ remember which $|_6$ one / of them / guest $|_7$ because it was $|_8$ so long ago.| (Non Local disjunction)

(15) What don't I remember?

i. Which of Bill and Sue it was. ii. Which guest it was.

Items were interspersed with 32 items from two unrelated experiments and 28 non-experimental filler items, and were counterbalanced and presented to subjects in an individually randomized order.

The manipulation was designed to explore whether violations of Locality and Alternatives on Demand influence online sentence interpretation, and whether the time course of the effect is modulated by the content of the inner restrictor. Assuming that remnants with fewer competitors will locate the correlate faster during retrieval, we expect that processing costs will

appear at earlier time points for restrictors with strong diagnostic content (Plural, Nominal) than for those with weak diagnostic content (Singular). In other words, we predict that the timing of any penalty associated with violating Locality and Alternatives on Demand will be modulated by how well the cues in the restrictor of the remnant distinguish the correlate from its competitors: restrictors with strongly diagnostic cues will speed retrieval of the correlate, thereby showing earlier violations, than restrictors which present weakly diagnostic cues for retrieval.

3.1.1 Interpretation question data

The percentage of Disjunction-antecedent responses was computed for each condition and is presented in Table 3, along with standard errors. As the Disjunction-antecedent responses for the Nominal condition were the result of error, responses for this condition were transformed to reflect the percent correct response, resulting in a much better model fit. A linear mixed effects logistic regression model was computed with Locality, Remnant type, and their interaction as fixed effect predictors; Table 4. Local correlates and Nominal restrictors were treated as the baseline in sum-coded contrasts. Random slopes and intercepts were computed for by-subject and by-item effects.

Remnant type	Local disjunction	Non-Local disjunction	Difference: Locality bias
Nominal	5% (1)	3% (1)	2%
Plural	98% (1)	70% (3)	28%
Singular	88% (3)	31% (2)	57%

Table 3. Experiment 2A: Mean percent of Disjunction-antecedent responses by condition. Standard errors are in parentheses.

	Estimate	Std. Error	Wald *Z*	*p*-estimate
(Intercept)	-0.12	0.18	-0.69	0.49
Locality	0.98	0.13	7.79	< 0.001

Plural	2.70	0.20	13.71	< 0.001
Singular	0.77	0.14	5.33	< 0.001
Plural x Locality	0.62	0.19	3.36	< 0.001
Singular x Locality	0.60	0.15	4.15	< 0.001

Table 4. Experiment 2A. Fixed effects from a logistic linear mixed-effects regression model on correct responses.

In general, there were more Disjunct-antecedent responses for Singular pronouns (M = 59%, SE = 2), z = 5.33, and Plural pronouns (M = 84%, SE = 2), z = 13.71, than the Nominal restrictor baseline (M = 4%, SE = 2). There was also an overall effect of Locality: Local nouns (M = 63%, SE = 2) were more likely to be selected as correlates than Non Local ones (M = 35%, SE = 2), z = 7.79. However, the effect of Locality was moderated by interactions with the Pronoun continuations: whether the Disjunct was in a Local position had a greater effect for Singular pronouns (*diff* = 57%), z = 4.15, and Plural pronouns (*diff* = 28%), z = 3.36, compared to Nominal restrictor conditions (*diff* = 2%). This result adds another dimension to the main pattern elicited in Experiment 1: disjunctions garnered a greater proportion of Disjunction-antecedent responses when in object position than when in subject position, but the effect holds most strongly for remnants with weakly diagnostic pronominal restrictors. In all, the interpretation results are very similar to those reported in the previous offline experiments.

3.1.2 Reading time data

On the reading time data, observations for each condition at each region were winsorized (Dixon, 1960; Tukey, 1962), so that scores above the 95th or below the 5th percentile were censored to the score at the 95th and 5th percentile, respectively. Means and standard errors were collected for all conditions at every region; however, as the analyses below only address Regions 6 to 8, the mean reading times and standard errors for those regions are presented in Table 5.

Figure 1 provides normalized reading times for regions of interest. As linear mixed effect regression models with random slopes and intercepts failed to converge, models with only random intercept models are presented in Table 6.

Restrictor type	Region 6 (Restrictor)		Region 7 (Spill over)		Region 8 (Final)	
	Local	Non-Local	Local	Non-Local	Local	Non-Local
Nominal	496 (9)	552 (12)	485 (8)	522 (11)	552 (11)	551 (11)
Plural	516 (11)	523 (10)	545 (12)	603 (14)	621 (13)	756 (21)
Singular	486 (8)	466 (7)	553 (12)	565 (12)	769 (22)	762 (21)

Table 5. Experiment 2A: Reading times in ms for regions of interest with standard errors in parentheses.

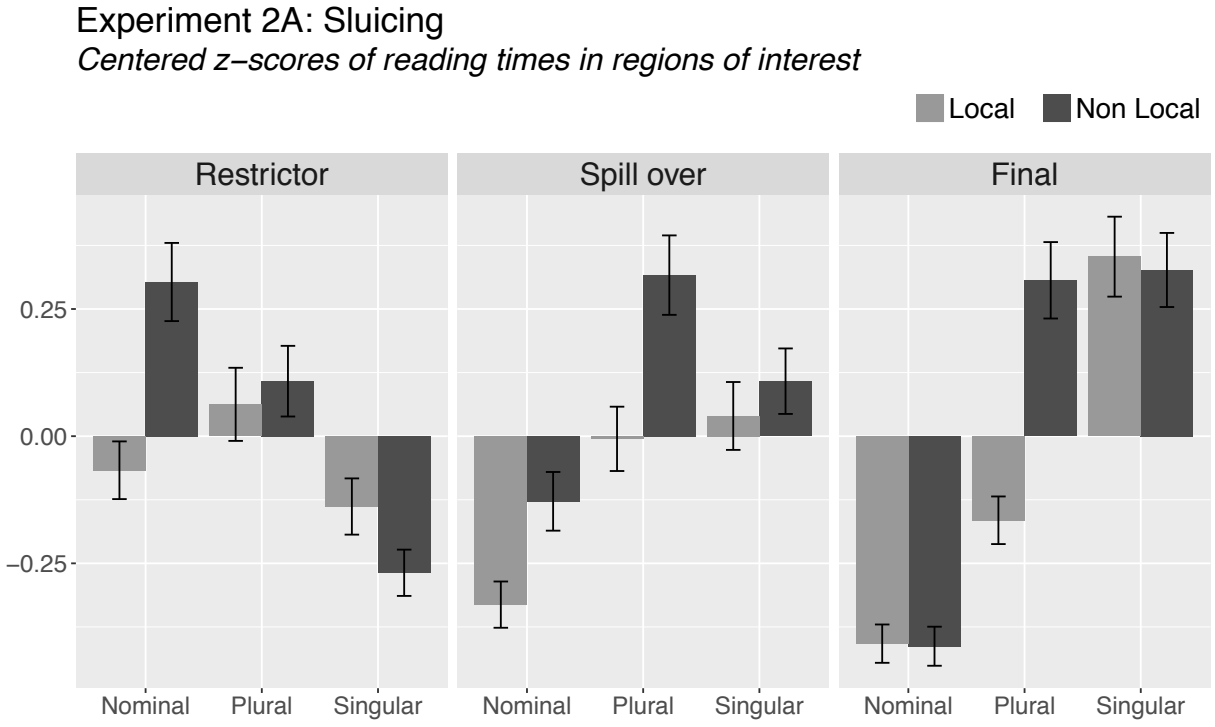

Fig 1. Experiment 2A: Centered z-scores of reading times in regions of interest. Locality effects were observed for Nominal restrictors in Restrictor and Spill over regions, and for Plural restrictors in Spill over and Final regions.

On the restrictor region, there was a general 15ms penalty for items with Non Local correlates compared to those with Local correlates, $t = 2.26$. However, the effect appeared only in the Nominal restrictor condition in planned paired by-subject and by-items t-test comparisons, with Bonferroni corrections [$t_1(47) = 4.79$, $p < 0.001$; $t_2(29) = 3.58$, $p < 0.01$]. Items with Singular pronouns showed a reading time advantage ($M = 476$, $SE = 5$) over the Nominal restrictors ($M = 524$, $SE = 7$), $t = -2.65$. Although Non Local correlates showed a processing advantage over Local correlates for Singular pronoun continuations (*diff* = 20ms) in the regression model, the difference was not significant in paired t-test comparisons.

In the spill over region, the Nominal restrictor baseline ($M = 504$, $SE = 7$) was read faster than Singular pronoun ($M = 559$, $SE = 8$), $t = 2.56$, and Plural pronoun ($M = 574$, $SE = 9$), $t = 5.40$, conditions. Items with Non Local correlates elicited a 36ms penalty over those with Local correlates, $t = 4.85$, which significant or near significant in both the Nominal restrictor [$t_1(47) = 3.36$, $p < 0.01$; $t_2(29) = 2.32$, $p = 0.06$] and Plural pronoun [$t_1(47) = 6.21$, $p < 0.001$; $t_2(29) = 6.25$, $p < 0.001$] conditions. In addition, there were two interactions between Pronoun restrictor and Locality: whereas Plural continuations showed a greater penalty for violating Locality than Nominal controls did, $t = 2.15$, Singular pronouns showed a reduced effect, $t = -2.24$.

On the sentence-final region, there was again an overall effect of Locality, manifesting in a 41ms slow down, $t = 3.44$. There were two interactions moderating this effect: Plural restrictors showed a greater effect of Locality, $t = 5.38$, whereas Singular pronouns again elicited a smaller penalty, $t = -2.86$. Only in the Plural restrictor conditions did Non Local disjunctions elicit a processing slow down in planned by-subjects and by-items t-test comparisons with Bonferroni corrections [$t_1(47) = 6.21$, $p < 0.001$; $t_2(29) = 6.25$, $p < 0.001$].

Responses from the interpretation questions were included as a factor in a post-hoc analysis to determine the effect that the post-sentence interpretation had on reading time within the sentence, depicted in Figure 2. The resulting models were a better fit than the planned models presented above for the spill over and final regions. The primary difference in these models was a three-way interaction between Locality, Remnant type, and Interpretation. For the spill over region, Non Local disjunctions resulted in a 42ms advantage for Singular pronouns when readers selected an Indefinite as the antecedent, but a 5ms penalty when they selected the Disjunct as the correlate [$\widehat{\beta} = 48.90$, $SE = 20.07$, $t = 2.436$, $p < 0.05$]. For the final region, Non Local disjunctions elicited an 84ms advantage for Singular pronouns when readers took an indefinite as the correlate, but an 80ms penalty when a Non Local disjunction was chosen as the antecedent [$\widehat{\beta} = 69.56$, $SE = 32.76$, $t = 2.123$, $p < 0.05$]. In essence, if the noun that readers later selected as the correlate was in Non Local position, reading times slowed. The effect was especially strong for Non Local disjunctive correlates. Notably, the dependence on interpretation was observed only for weakly diagnostic Singular pronouns, and not for other types of content in the restrictor.

	Region 6 (Restrictor)			Region 7 (Spill over)			Region 8 (Final)		
	Estimate	SE	*t*-value	Estimate	SE	*t*-value	Estimate	SE	*t*-value
(Intercept)	489.09	23.35	20.95*	545.68	17.02	32.06*	668.5	21.84	30.6*
Non Local	7.19	3.18	2.26*	18.02	3.72	4.85*	20.91	6.08	3.44*
Singular	-23.77	8.98	-2.65*	13.47	5.25	2.56*	96.80	8.60	11.25*
Plural	9.50	5.88	0.11	28.36	5.25	5.40*	20.10	8.60	2.34*
Length	2.59	2.89	0.37	--	--	--	--	--	--
Non Local x Singular	-17.03	4.50	-3.79*	-11.79	5.25	-2.24*	-24.64	8.60	-2.86*
Non Local x Plural	-3.75	4.50	0.41	11.31	5.25	2.15*	46.28	8.60	5.38*

Table 6. Experiment 2A: Fixed effects in linear mixed regression models for regions of interest. '*' marks *t*-values that were considered fully significant, at levels greater than |2|.

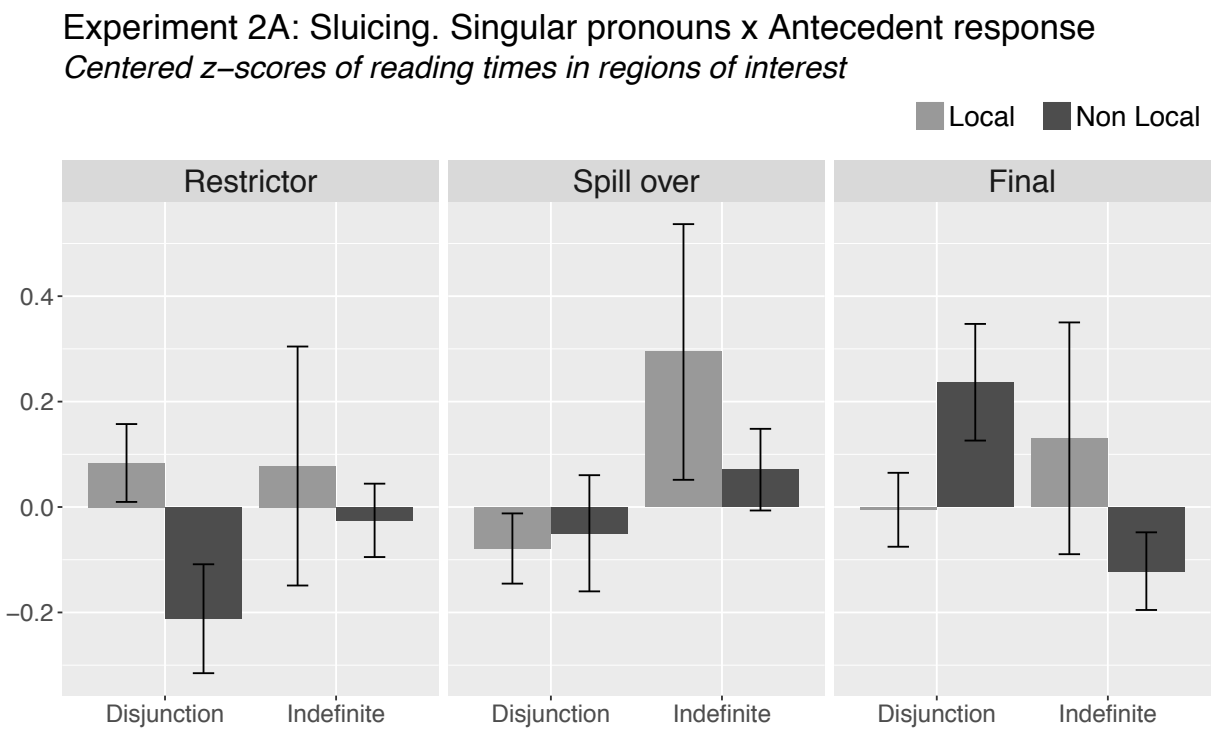

Fig 2. Experiment 2A: Effect of post-sentence interpretation on normalized reading times for regions of interest for Singular restrictors.

3.1.3 Discussion

Results showed a clear reading time penalty for sluices that violated the Locality Bias (Frazier & Clifton, 1998, among others). However, the time course of the penalty was affected by the restrictor type. In line with Harris (2015), there was an early effect of Locality for Nominal restrictors on the remnant region. However, the Locality penalty for Nominal restrictors dissipated by the following region and was supplanted by a general advantage in the final region. For Plural restrictors, the Locality effect appeared after the remnant for the remainder of the sentence. However, Locality did not affect reading times on Singular restrictors. The expected effect was observed on the final region only when interpretation was also taken into consideration. As Singular items contained only weakly diagnostic cues, subjects may have

delayed resolving the correlate-remnant relation until they had read as much of the sentence as they could.

The materials presented in the experiments so far have all been instances of standard sluicing ellipsis. We now turn to similar constructions containing a clefted *wh*-element (16b). In these cases, the *which N* phrase can be analyzed as the pivot of the cleft *it was*, so that the embedded clause is not syntactically isomorphic to the antecedent. I will call such cases "sluicing clefts" as a cover term, without adopting a detailed analysis. The underlying syntax of the it-cleft has two central analyses in the literature: either the pivot is base-generated in its surface position, or it is moved to from a position within an embedded clause (see Haegeman, Meinunger, & Vercauteren, 2015, for review). In either case, I will assume, for concreteness, that they are formed by eliding the embedded clause *that left* (16b), parallel to ellipsis of the IP in standard sluicing (16a).

(16) a. *Standard sluice:* Someone left, but I don't know who <left>

b. *Sluicing cleft:* Someone left, but I don't know who it was <that left>

The relation between sluicing clefts (16b) and more standard cases of sluicing (16a) is a currently matter of debate (e.g., Vicente, to appear, for review). On the one hand, sluicing with cleft sources might be only superficially related to true sluicing, and present an ellipsis source that is preferred in some languages but not others (e.g., Mechant's 1998 pseudosluicing account for Japanese, or van Craenenbroek's (2010) analysis of spading in Frisian). On the other, sluicing of either variety might represent fundamentally the same phenomenon: although they have different underlying structural sources of ellipsis, their semantic equivalence licenses the ellipsis

(e.g., Barros, 2014). For our purposes, the important point is that it-cleft variants have been argued to be particularly natural with antecedent clauses containing disjunctions (AnderBois, 2011; Barros, 2014). In addition, they appear to impose a uniqueness requirement on the correlate; (16b), for example, indicates that only one individual left (Barros, 2014).

It is theoretically possible that previous materials were ambiguous between a standard monoclausal structure, and a biclausal cleft structure, and that the choice of preferred correlate might depend on which structure was selected in some way. Although the basic processing predictions remain the same between standard sluices and sluicing clefts, the materials are disambiguated towards the cleft structure in the studies below. Examining sluicing clefts also allows us to address the extent to which the Locality Bias should be attributed to a general pressure for parallelism between the antecedent and the ellipsis clause. If Locality reflects syntactic parallelism, then sluicing clefts should show a reduced effect of Locality, as the antecedent and the elided clause cannot be syntactically isomorphic, and therefore must be non-parallel in some sense. However, if the sluicing clefts show similar processing patterns as standard sluicing, Locality is more likely to reflect a correlate-remnant pairing process, along the lines of the second task described in (4).

3.2 Experiment 2B

A distinct set of forty-eight native speakers of English from the Claremont Colleges participated in the study for $10. The materials and method for this follow up experiment was identical to Experiment 2A, except that an it-cleft (*it was*) phrase was included to avoid ambiguity (17). Sentences were followed by the same comprehension questions as in the previous study (15).

(17) a. $|_1$ A guest $|_2$ talked to $|_3$ Bill or Sue, $|_4$ but I don't $|_5$ remember which $|_6$ one / of them / guest $|_7$ it was $|_8$ at this point.| (Local disjunction)

b. $|_1$ Bill or Sue $|_2$ talked to $|_3$ a guest, $|_4$ but I don't $|_5$ remember which $|_6$ one / of them / guest $|_7$ it was $|_8$ at this point.| (Non Local disjunction)

Items were interspersed with 52 items from two unrelated experiments and 28 non-experimental filler items, and presented to subjects in individually randomized counterbalanced order.

3.2.1 Interpretation question data

The interpretation question data was modeled as in the previous study. Means and standard errors are provided in Table 7, and the logistic linear mixed effects regression model in Table 8.

Remnant type	Local disjunction	Non-Local disjunction	Difference: Locality bias
Nominal	2% (1)	3% (1)	1%
Plural	97% (1)	91% (2)	6%
Singular	78% (3)	49% (3)	29%

Table 7. Experiment 2B: Mean percent of Disjunction-antecedent responses by condition. Standard errors are in parentheses.

	Estimate	Std. Error	Wald *Z*	*p* estimate
(Intercept)	2.478	0.192	12.936	< 0.001
Locality	-0.499	0.117	-4.247	< 0.001
Plural	0.811	0.183	4.440	< 0.001
Singular	-1.724	0.141	-12.263	< 0.001
Plural x Locality	-0.161	0.182	-0.883	0.377
Singular x Locality	-0.285	0.135	-2.111	< 0.05

Table 8. Experiment 2B. Fixed effects from a logistic linear mixed-effects regression model on correct responses.

First, there was again an overall effect of Locality: items obeying the Locality Bias (M = 91%, SE = 1) were more likely to receive a "correct" response (the Disjunction for pronominal restrictors; the Indefinite for nominal restrictors) than Non local conditions (M = 79%, SE = 2), z = -4.25, $p < 0.001$, regardless of Remnant type. Plural restrictors elicited more Disjunction-antecedent responses (M = 94%, SE = 1) compared to the grand mean (M = 85%, SE = 1), z = 4.44, $p < 0.001$, whereas the ambiguous Singular restrictors exhibited far fewer Disjunction-antecedent response (M = 64%, SE =2), z = -12.26, $p < 0.001$. Crucially, the likelihood of a correct Disjunction-antecedent response was differentially increased for Singular continuation when the Locality Bias was satisfied ($diff$ = 30%) over Indefinite ($diff$ = 1%) continuations, z = -2.11, $p < 0.05$. In contrast, Plural ($diff$ = 6%) continuations were not affected by Locality, $z < 1$. Plural and Nominal restrictors provided cues that effectively distinguished the target correlate from its competitor. The difference between Local and Non Local correlates was less dramatic for the strongly diagnostic Plural pronouns in sluicing clefts compared to the sluiced structures in the previous study (6% vs. 28%). It is unclear whether any structural difference between the two would have motivated this difference, since the interpretation of Singular pronoun conditions in this experiment is closer to offline results of true sluicing, reported in Experiment 1.

3.2.2 Reading time data

Prior to analysis, reading time scores were winsorized as in the previous experiment; see Table 9 and Figure 3. Maximal random effects models are presented in Table 10.

Restrictor type	Region 6 (Restrictor)		Region 7 (Spill over)		Region 8 (Final)	
	Local	Non Local	Local	Non Local	Local	Non Local
Nominal	488 (10)	490 (10)	418 (11)	450 (9)	525 (15)	521 (14)
Plural	467 (9)	495 (11)	429 (11)	472 (11)	658 (24)	761 (30)

Singular	439 (6)	431 (6)	428 (10)	439 (9)	743 (26)	929 (44)

Table 9. Experiment 2B: Reading times in ms for regions of interest with standard errors in parentheses.

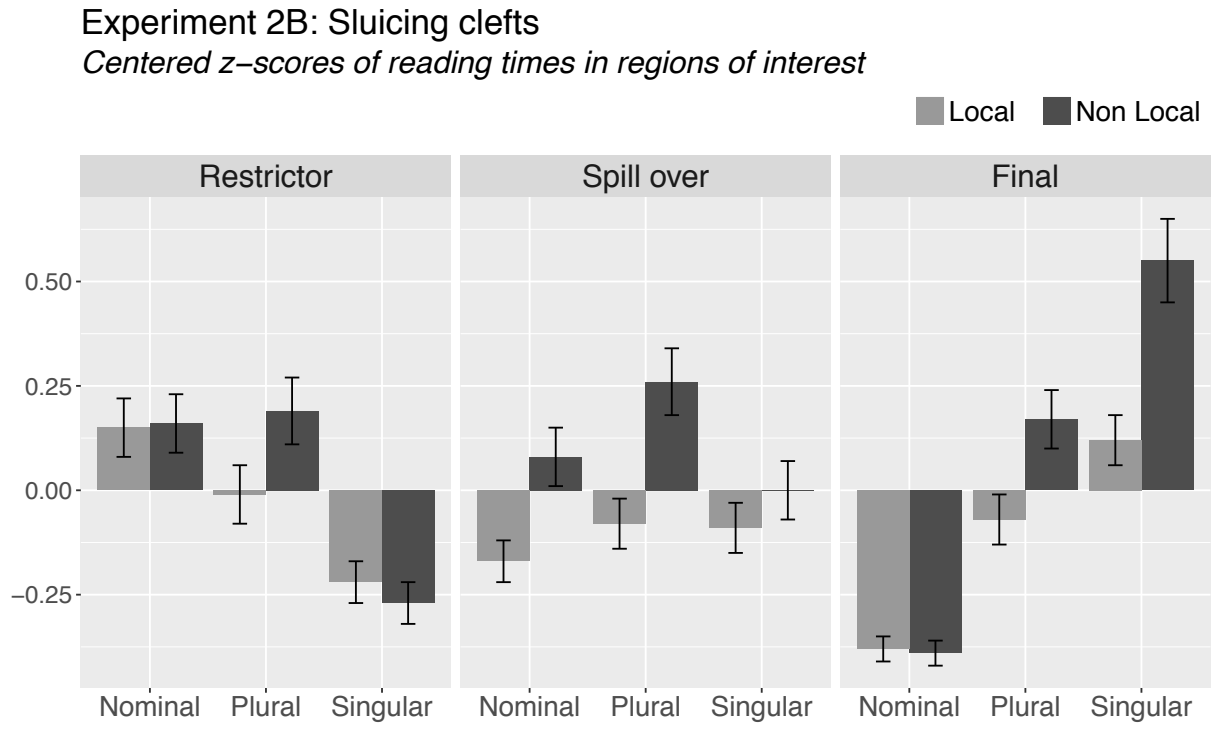

Fig 3. Experiment 2B: Centered z-scores of reading times in regions of interest. Locality effects were observed for Plural restrictors in all regions, for Nominal restrictors in the Spill over region, and for Singular restrictors in the Final region.

On the region containing the restrictor (Region 6), Singular pronoun restrictors elicited faster overall reading times (M = 435, SE = 4) compared to the Nominal baseline (M = 489, SE = 7), t = -3.01, but there was no main effect for Plural restrictors (M = 481, SE = 7). However, Plural restrictors showed a greater penalty for violating Locality (*diff* = 28) compared to Nominal restrictors (*diff* = 2), t = 2.95. The Locality difference for Plural restrictors was significant in paired by-subjects and by-items t-tests with Bonferroni corrections. Surprisingly, Singular restrictors showed a mild advantage for Non Local disjunctions (*diff* = 8ms), t = -2.35, but this

effect was most likely spurious, as the effect was not significant in either by-subjects or by-items planned t-test comparisons.

In the spillover region (Region 7), numerous effects were observed. As predicted, structures that violated the Locality Bias were costly, eliciting slower reading times (M = 454, SE = 5) in comparison to structures that did not (M = 425, SE = 4), t = 5.66. Plural restrictors not only elicited slower reading times overall, but they were also more greatly penalized (*diff* = 43) when violating Locality than Nominal counterparts (*diff* = 32), t = 2.29. Both Plural and Nominal restrictors elicited reading time costs for Non Local correlates in by-subjects and by-items t-tests with Bonferroni corrections [Plural: $t_1(47) = 4.47$, $p < 0.001$; $t_2(29) = 3.15$, $p < 0.05$; Nominal: $t_1(47) = 3.74$, $p < 0.001$; $t_2(29) = 3.34$, $p < 0.05$]. Singular continuations again appeared to be insensitive to Locality at this region: although there was an effect of Locality in the regression model, the differences were not detected in planned t-test comparisons.

In the sentence final region (Region 8), violating Locality resulted in a 95ms reading time penalty, t = 4.87. Singular pronoun restrictors elicited delayed reading times (M = 836, SE = 26) over the Nominal restrictor baseline (M = 580, SE = 17), t = 5.35, in general, and were differentially penalized for violating Locality (*diff* = 186) over Nominal (*diff* = 4) counterparts, t = 2.96. The penalty for Non Local correlates to remnants with Singular pronoun restrictors was observed in planned by-subjects and by-items t-tests [$t_1(47) = 3.19$, $p < 0.01$; $t_2(29) = 3.81$, $p < 0.01$]. However, the comparable interaction did not manifest for Plural continuations (*diff* = 103), despite a clear numerical trend for Locality which was significant in both by-subjects and by-items t-test comparisons [$t_1(47) = 3.72$, $p < 0.01$; $t_2(29) = 2.73$, $p < 0.05$].

	Region 6 (Restrictor)			Region 7 (Spill over)			Region 8 (Final)		
	Estimate	SE	*t*-value	Estimate	SE	*t*-value	Estimate	SE	*t*-value
(Intercept)	446.17	19.88	22.45*	439.29	11.89	36.94*	641.17	31.24	20.52*

Non Local	3.99	2.69	1.48	14.74	2.61	5.66*	96.33	19.77	4.87*
Singular	-23.38	7.77	-3.01*	-5.89	3.68	-1.59	105.80	19.78	5.35*
Plural	6.93	5.03	1.38	10.56	3.68	2.87*	12.08	19.77	0.61
Length	3.94	2.52	1.57	--	--	--	--	--	--
Non Local x Singular	-8.94	3.81	-2.35*	-10.76	3.69	-2.92*	82.74	27.99	2.96*
Non Local x Plural	11.21	3.80	2.95*	8.45	3.69	2.29*	14.26	27.96	0.51

Table 10. Experiment 2B: Fixed effects in linear mixed regression models for regions of interest. '*' marks *t*-values that were considered fully significant, at levels greater than |2|.

3.2.3 Discussion

The overall effects for sluicing clefts in Experiment 2B were largely comparable to the sluices explored in Experiment 2A, suggesting that the Locality Bias occurs regardless of the source of ellipsis. Similar general effects of Locality were observed in the two experiments, albeit on slightly later regions, and were modulated by the strength of the cue provided in the restrictor. Most importantly, Singular restrictors showed a delayed effect for Locality compared to the Plural restrictors. This delay is expected if readers sometimes defer making anaphoric inferences when they are not presented with enough information to determine the referent, possibly opting for a shallower, underspecified representation initially (Greene, McKoon, & Ratcliff, 1992; Levine, Guzmán, & Klin, 2000; Klin et al., 2006; Stewart, Holler, & Kidd, 2007). Readers might engage in a 'wait and see' strategy for pronoun resolution: postponing an anaphoric inference potentially allows readers to accumulate more information with which to make a more informed decision. Interestingly, the Locality effect appeared in the final region for Singular restrictors even without considering the post-sentential resolution of the remnant.

The following experiment manipulates whether prior contexts support situations in which the indefinite noun is likely to be one of multiple entities, e.g., a guest at a party. On the one hand, if retrieving a correlate for an ambiguous remnant is governed by strategic delay due to

lack of information, contexts that specify multiple entities should facilitate retrieval of the indefinite as the correlate for a remnant. On the other, if the process is sensitive only to the form of the antecedent clause, Locality preferences guiding retrieval might be unaffected by prior context.

4 Experiment 3

The following self-paced reading experiment investigates whether the processing delay for Singular pronominal restrictors is modulated by prior context biasing towards a multiple-individual interpretation of the indefinite in sluicing clefts similar to Experiment 2B. Thirty-two native speakers of English from the Claremont Colleges participated in the experiment for $10 each. Materials consisted of 20 sentences with singular pronoun (*which one*) continuations from Experiment 1, which were given contexts (18), provided in Appendix C. The experimental manipulation crossed Locality of the disjunction (*Local* vs. *Non Local*) and Context (*Neutral* vs. *Biased to Indefinite*). Participants read contexts and target sentences, presented on different lines, at their own pace in a self-paced reading task. After completing the sentence, they answered an interpretation question about the target sentence (19).

(18) *Neutral context:* It was a particularly humid night.

Biased context: The party was swarmed with people.

a. $|_1$ A guest $|_2$ talked to $|_3$ Bill or Sue, $|_4$ but I don't remember $|_5$ which one $|_6$ it was $|_7$ at the moment.

b. $|_1$ Bill or Sue $|_2$ talked to $|_3$ a guest, $|_4$ but I don't remember $|_5$ which one $|_6$ it was $|_7$ at the moment.

(19) *Interpretation question:* What don't I remember?

i. Which of Bill and Sue it was. ii. Which guest it was.

Items were interspersed with 56 items from three unrelated experiments and 34 non-experimental filler items, and presented to subjects in individually randomized counterbalanced order. After the reading block, participants completed an exit questionnaire with questions about the context sentences, as in *Does the sentence suggest there were multiple guests?*. Participants selected one response from three possible answers: *Yes*, *No*, and *Not sure*. The third response was necessary, as the question would have sometimes seemed unnatural in neutral contexts.

4.1 Interpretation data

Results from the exit questionnaire are provided in Table 11. There are two natural ways to group the data, and both converge on the same overall result. First, comparing affirmative (*Yes*) against non-affirmative (*No/Not sure*) responses, biased contexts (M = 86%, SE = 2) were more readily associated with multiple entities than neutral contexts (M = 10%, SE = 2), z = 10.30, $p < 0.001$. Second, if negative (*No*) responses are compared with non-negative (*Yes/Not sure*), biased contexts (M = 93%, SE = 1) were again judged to support multiple entities than neutral ones (M = 44%, SE = 3), z = 7.05, $p < 0.001$. In short, putatively biased contexts may be assumed to have effectively biased interpretation towards multiple entities.

Response	*Context*	
	Bias	Neutral
Yes	86%	10%
No	7%	56%
Not sure	7%	34%
Grouping method 1 (*Yes only*)	86%	10%

Grouping method 2 (*Yes / Not sure*)	93%	44%

Table 11. Experiment 3: Efficacy of contexts indicating multiple entities for the indefinite. The mean percentage of each response is shown, along with two methods for grouping *Not sure* responses.

With respect to responses to interpretation questions (19), Disjunction-antecedent responses were more greatly preferred when in object (M = 94%, SE = 1) than subject (M = 69%, SE = 3) position, z = -7.73, $p < 0.001$, but context did not affect interpretation; see Table 12.

	Estimate	Std. Error	Wald Z	p estimate
(Intercept)	2.199	0.300	7.328	< 0.001
Biased context	-0.028	0.141	-0.200	0.842
Non Local	-1.155	0.149	-7.734	< 0.001
Biased x Non Local	0.044	0.141	0.308	0.758

Table 12. Experiment 3: Fixed effects from logistic linear mixed-effects regression model on Disjunction-antecedent responses.

4.2 Reading time data

The region-by-region reading time was winsorized as before, and the means and standard errors for regions of interest are presented in Table 13; see also Figure 4. The results of the statistical model are presented in Table 14. Biased contexts elicited slower reading times in both the remnant (a 20ms cost), and spill over (a 19ms cost) region. No other effects were observed in these regions. However, a penalty for violating Locality was observed in the final region: Non Local disjunctions elicited longer reading times (M = 679, SE = 19) compared to Local disjunctions (M = 600, SE = 13), t = 4.05, $p < 0.001$. In addition, there was an interaction, so that violating Locality was greater for Biased (*diff* = 122) than Neutral (*diff* = 34) contexts, t = 2.25, $p < 0.05$. A model with by-item response bias from the exit questionnaire as a predictor was created for each region, but it did not yield materially different results.

	Region 6 (Remnant)		Region 7 (Spill over)		Region 8 (Final)	
Context	Local	Non-Local	Local	Non-Local	Local	Non-Local
Bias	491 (10)	497 (10)	460 (9)	477 (10)	591 (17)	713 (29)
Neutral	482 (9)	466 (8)	451 (9)	446 (8)	610 (19)	644 (25)

Table 13. Experiment 3: Reading times (ms) for regions of interest with standard errors in parentheses.

	Region 6 (Remnant)			Region 7 (Spill over)			Region 8 (Final)		
	Estimate	SE	*t*-value	Estimate	SE	*t*-value	Estimate	SE	*t*-value
(Intercept)	483.99	13.91	34.79*	458.34	13.59	33.73*	639.52	29.64	21.57*
Biased context	9.95	3.64	2.74*	9.79	3.40	2.88*	12.44	9.70	1.28
Non Local	-2.62	3.64	-0.72	3.03	3.40	0.89	39.23	9.70	4.05*
Biased x Non Local	5.64	3.64	1.55	5.44	3.40	1.60	21.84	9.70	2.25*

Table 14. Experiment 3: Fixed effects in linear mixed regression models for regions of interest. '*' marks *t*-values that were considered fully significant, at levels greater than |2|.

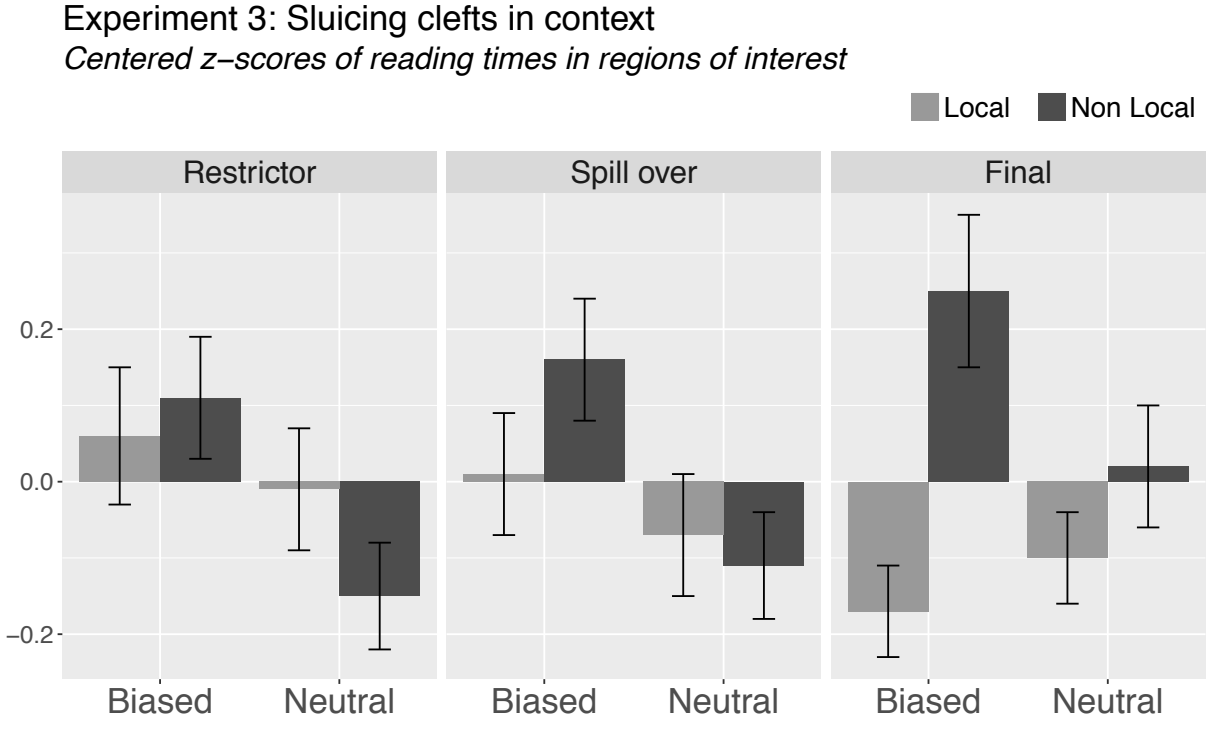

Fig 4. Experiment 3: Centered z-scores of reading times in regions of interest. A Locality effect was observed in the Final region, and was greater for Biased than for Neutral contexts.

4.3 Discussion

Experiment 3 replicated the results for the Singular pronominal restrictors from the previous experiments: Non Local disjunctions elicited longer reading times than Local disjunctions in the sentence final region. Further, contexts that biased towards multiple entity interpretations of the indefinite were found to increase reading times on the remnant and the spill over region, as well as to induce larger penalties for violating Locality in the sentence final region. The results also suggest that Biasing contexts were not sufficient to override Alternatives on Demand during the remnant-correlate pairing procedure (step 2 described in (4)), but instead made indefinite nouns into more tempting correlates, at least temporarily. The fact that context did not influence the ultimate interpretation of the sluice further indicates that readers eventually resolved the correlate to the disjunction, despite contextual support for the multiple entities interpretation of the indefinite.

5 General discussion

Three experiments, consisting of several questionnaire tasks and three online self-paced reading studies, probed the relationship between the Locality Bias, a structural economy preference favoring linearly or structurally more accessible antecedents in clausal ellipsis and related structures, and Alternatives on Demand, a discourse principle militating against computing discourse new alternatives. There was consistent evidence that the two constraints in question interact, not only in offline interpretation, but also in real time sentence processing.

The results further support a model in which the sentence position of possible correlates strongly affects how remnants to sluicing and sluicing cleft structures are resolved, and that the time course of this influence is determined, at least in part, by the content of the restrictor within

the remnant. Although remnants with both strong and weakly diagnostic cues showed an online penalty for violating the Locality Bias, the effect was delayed for remnants with weakly diagnostic content. One possible explanation is that readers strategically left the pairing between the remnant and the correlate underdetermined until the end of the sentence. Another possibility is that readers were able to pair the remnant and the correlate faster when they had sufficient information to begin the retrieval process: plural pronoun and nominal restrictors provided more restrictive or diagnostic cues, which allowed the processor to retrieve the correct correlate at an earlier time point. This pattern could be fruitfully modeled with current cue-based approaches to sentence parsing (e.g., Lewis et al., 2006), in which retrieval of an item is subject to interference from similar items. Taking the restrictor content as a kind of cue for pairing the remnant with the correlate, weakly diagnostic pronominal cues (*one*) may be subject to more interference from multiple correlate targets, as no feature sufficiently distinguishes one potential correlate from another. In contrast, strongly diagnostic cues (unambiguous pronoun or nominal restrictors) specify the target more directly, thereby allowing violations of Locality to be detected more quickly (Harris, 2015).

In addition, the results suggest that the likelihood of a constituent serving as a correlate is influenced by its discourse context during sentence processing. Contexts biasing towards a multiple entity interpretation for indefinites were associated with larger penalties for violating the Locality Bias, suggesting that indefinites became more viable competitors in the correlate-remnant pairing or retrieval process. Yet, most current theories of cue-based parsing discuss the retrieval mechanism primarily in terms of surface or morphological features that are compared via an automatic parallel cue matching procedure. An alternative possibility is that at least some types of retrieval access richer types of information, such as a discourse or conceptual level

representations, in addition to surface or morphological cues, a possibility which has been discussed in connection with pronouns more generally (e.g., Cloitre & Bever, 1988; Koh & Clifton, 2002).

In addition, the general pattern held for both cases of standard sluicing and sluicing clefts, raising the possibility that that the locus of the effect should be attributed to general processes governing the resolution of the correlate, rather than to the process of reconstructing an isomorphic syntactic structure at the ellipsis site. In general, the basic processing tasks required to interpret clausal ellipsis articulated in (4) presuppose an interdependence between pairing the remnant with a correlate, on the one hand, and interpreting the ellipsis, on the other. More research is required to determine whether one truly precedes the other, and the extent to which general parsing preferences guiding ellipsis resolution, such as parallelism between clauses, should be attributed to one or more of these processes.

The results also motivate a theoretical advance beyond a broad notion of discourse salience, which does not suffice to capture the results obtained here. The principal reason is that the first-mentioned noun is usually thought of as the most salient entity in the sentence, and is understood as the sentence topic (e.g., Reinhart, 1981). If we tried to capture the effect of the disjunction solely in terms of increased salience instead of discourse alternatives, we would expect more Disjunction-antecedent responses for subject-position correlates, not fewer, as the salience of disjunctions should theoretically increase when occupying subject position. This is not to say that salience does not have a role in finding a correlate for the remnant position, only that it would be difficult to explain the results exclusively in terms of the notion. In addition, a general preference to interpret a subject-position indefinite noun as the sentence topic may have

amplified the Locality effect for object-position disjunctions, as the indefinite would be an even less likely candidate for a multiple entities construal.

Overall, the results pose a more general challenge to current accounts of linguistic dependency formation in exposing an explanatory gap between how processing heuristics interact to create online representations and how such representations are utilized within routines guiding retrieval. However that gap is filled, it seems relatively clear that constraints governing discourse interpretation provide an important source of information that exerts a powerful influence at even the earliest stages of retrieval and sentence comprehension.

Acknowledgements

Although a simple "thank you" fails to convey the gratitude I feel towards Lyn Frazier for years of mentorship and support, it will have to do. I also thank Joel Fishbein for administering Experiments 2 and 3. I am indebted to Chuck Clifton and Lyn Frazier for discussion on methodology and interpretation, to the UCSC Department of Linguistics for feedback during a colloquium talk, to participants at Architectures and Mechanisms for Language Processing held at Riva del Garda, Italy, participants in the UCLA Psycholinguistics Seminar, the 26th Annual CUNY Sentence Processing Conference hosted by the University of South Carolina, and to the audience at LynSchrift 2018, where portions of this work were previously presented. I thank UCLA and Pomona College for financial support. Any errors are naturally my own.

References

AnderBois, S. (2014). The semantics of sluicing: Beyond truth conditions. *Language*, *90*, 887–926. doi: 10.1353/lan.2014.0110

Alonso-Ovalle, L. (2006). *Disjunction in alternative semantics*. Ph.D. thesis, UMass Amherst.

Arnold, J. E. (1998). *Reference form and discourse patterns.* Ph.D. thesis, Stanford University.

Arnold, J. E. (2010). How speakers refer: the role of accessibility. *Language and Linguistics Compass*, 4, 187–203. doi: 10.1111/j.1749-818x.2010.00193.x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.

Barr, D. J, Levy, R. Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. doi: 10.1016/j.jml.2007.12.005

Barros, M., & Vicente, L. (2016). A remnant condition for ellipsis. In: *Proceedings of WCCFL* (Vol. 33, pp. 57-66).

Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31.

Beaver, D., & Clark, B. Z. (2008). *Sense and Sensitivity: How Focus Determines Meaning.* Oxford: Wiley-Blackwell.

Carlson, K. (2002). *Parallelism and Prosody in the Processing of Ellipsis Sentences*. Outstanding Dissertations in Linguistics, New York, NY: Routledge.

Carlson, K., Dickey, M. W., Frazier, L., & Clifton, Jr., C. (2009). Information structure expectations in sentence comprehension. *Quarterly Journal of Experimental Psychology*, 62, 114–139. doi: 10.1080/17470210701880171

Chung, S., Ladusaw, W., & McCloskey, J. (1995). Sluicing and logical form. *Natural Language Semantics*, 3, 239–282. doi: 10.1007/bf01248819

Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry*, 24, 239–297.

Cloitre, M., & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes*, 3, 293–322. doi: 10.1080/01690968808402092

Dayal, V., & Schwarzschild, R. (2010). Definite inner antecedents and *wh*-correlates in sluices. *Rutgers Working Papers in Linguistics*, *3*, 92–114.

Dickey, M. W., & Bunger, A. C. (2011). Comprehension of elided structure: Evidence from sluicing. *Language and Cognitive Processes*, 26, 63–78. doi: 10.1080/01690961003691074

Dixon, W. J. (1960). Simplified Estimation from Censored Normal Samples. *Annals of Mathematical Statistics*, *31*, 385–391. doi: 10.1214/aoms/1177705900

Fraundorf, S. H., Benjamin, A. S., & Watson, D. G. (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language, 69*, 196–227. doi: 10.1016/j.jml.2013.06.003

Frazier, L. (1987). Sentence processing: A tutorial review. In M Coltheart (Ed.), *Attention and Performance XII*, (pp. 559–586). Hillsdale, NJ: Lawrence Erlbaum Associates.

Frazier, L, Carlson, K., & Clifton, Jr., C. (2006). Prosodic phrasing is central to language comprehension. *TRENDS in Cognitive Sciences*, 10, 244-249. doi: 10.1016/j.tics.2006.04.002

Frazier, L., & Clifton, Jr., C. (1998). Comprehension of sluiced sentences. *Language and Cognitive Processes*, 13, 499–520. doi: 10.1080/016909698386474

Frazier, L., & Clifton, Jr., C. (2001). Parsing coordinates and ellipsis: Copy α. *Syntax*, 4, 1–22. doi: 10.1111/1467-9612.00034

Frazier, L., & Clifton, Jr., C. (2002). Processing "d-linked" phrases. *Journal of Psycholinguistic Research*, 31, 633–659. doi: 10.1023/A:1021269122049

Frazier, L., & Clifton, Jr., C. (2005). The syntax-discourse divide: processing ellipsis. *Syntax*, 8, 121–174. doi: 10.1111/j.1467-9612.2005.00077.x

Frazier, L., & Clifton, Jr., C. (2011). D-linking and Memory Retrieval: The Annoying Case of Sluicing. In Harris, J. A., & Grant, M. (Eds.), *University of Massachusetts Occasional Papers in Linguistics: Processing Linguistic Structure*. Volume 38 (pp. 37–52). Amherst, MA: GLSA Publications.

Garnham, A., Oakhill, J., Ehrlich, M. F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, 34, 41–62. doi: 10.1006/jmla.1995.1003

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 266-283. doi: 10.1037/0278-7393.18.2.266

Haegeman, L., Meinunger, A., & Vercauteren, A. (2015). The syntax of it-clefts and the left periphery of the clause. In Shlonsky, U. (Ed.), *Beyond Functional Sequence: The Cartography of Syntactic Structures, Volume 10* (pp. 73-90). Oxford: Oxford University Press.

Harris, J. A. (2015). Structure modulates similarity-based interference in sluicing: An eye tracking study. *Frontiers of Psychology, 6.* doi: 10.3389/fpsyg.2015.01839

Harris, J. A., & Carlson, K. (2016). Keep it local (and final): Remnant preferences in 'let alone' ellipsis. *Quarterly Journal of Experimental Linguistics*, *69*, 1278-1301.

doi: 10.1080/17470218.2015.1062526

Harris, J. A., & Carlson, J. (in press). Information structure preferences in focus-sensitive ellipsis: How defaults persist. To appear in *Language & Speech*. doi: 10.1177/0023830917737110

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 447–456. doi: 10.1016/j.jml.2007.11.007

Kaiser, E., & Trueswell, J.C. (2004). The role of discourse context in the processing of a flexible word-order *Cognition, 94*, 113-147. doi: 10.1016/j.cognition.2004.01.002

Kim, C. S., Gunlogson, C., Tanenhaus, M. K., & Runner, J. T. (2015). Context-driven expectations about focus alternatives. *Cognition*, *139*, 28-49. doi: 10.1016/j.cognition.2015.02.009

Klin, C. M., Guzmán, A. E., Weingartner, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language*, *54*, 131-143. doi: 10.1016/j.jml.2005.09.001

Koh, S., & Clifton, C. (2002). Resolution of the antecedent of a plural pronoun: Ontological categories and predicate symmetry. *Journal of Memory and Language*, *46* 830-844. doi: 10.1006/jmla.2001.2829

Kratzer, A., & Shimoyama, J. (2002). Indeterminate pronouns: The view from Japanese. *The Proceedings of the 3rd Tokyo Conference on Psycholinguistics* (pp. 1–25). doi: 10.1007/978-3-319-10106-4_7

Levine, W. H., Guzmán, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language*, *43*, 594-617. doi: 10.1006/jmla.2000.2719

Martin, A. E. (2010). *Memory operations and structures in sentence comprehension: Evidence from Ellipsis.* Ph.D. thesis, New York University, New York, NY.

Martin, A. E. & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, 58, 879–906. doi: 10.1016/j.jml.2007.06.010

Martin, A. E. & McElree, B. (2011). Direct-access retrieval during sentence comprehension: Evidence from sluicing. *Journal of Memory and Language*, 64, 327–343. doi: 10.1016/j.jml.2010.12.006

Merchant, J. (1998). 'Pseudosluicing': Elliptical clefts in Japanese and English. *ZAS Working Papers in Linguistics*, *10*, 88-112.

Merchant, J. (2001). *The Syntax of Silence*. Oxford: Oxford University Press.

Paterson, K. B., Liversedge, S. P., Filik, R., Juhasz, B. J., White, S. J., & Rayner, K. (2007). Focus identification during sentence comprehension: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *60*, 1423–1445. doi: 10.1080/17470210601100563

Pesetsky, D. (1987). *Wh*-in-situ: Movement and unselective binding. In Reuland, E., & ter Meulen, A. (Eds.), *The Representation of (In)definiteness*, (pp. 98–129). Cambridge, MA: MIT Press.

Poirier, J., Wolfinger, K., Spellman, L. & Shapiro, L. P. (2010). The real-time processing of sluiced sentences. *Journal of Psycholinguistic Research*, *39*, 411–427. doi: 10.1007/s10936-010-9148-9

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, *27*, 53–94.

Romero, M. (1998). *Focus and reconstruction effects in* wh-*phrases*. Ph.D. thesis, UMass Amherst.

Rooth, M. (1985). *Association with Focus.* Ph.D. thesis, UMass Amherst.

Ross, J. R. (1969). Guess who. In Binnick, R. et al. (Eds.), *Proceedings of the 5th Annual Chicago Linguistic Society* (pp. 252–286), Chicago, IL. doi: 10.1093/acprof:oso/9780199645763.003.0002

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language*, *46*, 341–370. doi: 10.1006/jmla.2001.2812

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation between Sound and Structure.* Cambridge: MIT Press.

Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *The Quarterly Journal of Experimental Psychology*, *60*, 1680-1696. doi: 10.1080/17470210601160807

Tukey, J. W. (1962). The Future of Data Analysis. *Annals of Mathematical Statistics*, *33*, 1–67. doi: 10.1214/aoms/1177704711

Van Craenenbroeck, J. (2010). The Syntax of Ellipsis: Evidence from Dutch dialects. New York, NY: Oxford University Press.

Van Dyke, J. A., & Johns, C. L. (2010). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, *6*, 193–211. doi: 10.1002/lnc3.330

Vicente, L. (to appear). Sluicing and its subtypes. In *The Oxford Handbook of Ellipsis* In Temmerman, T. and van Craenenbroeck, J. (Eds.). Oxford: Oxford University Press.

Watkins, O. C., & Watkins, M. J. (1975). Build-up of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 442–452. doi: 10.1037/0278-7393.1.4.442

Wolter, L., Gorman, K.S., & Tanenhaus, M.K. (2011). Scalar reference, contrast and discourse: Separating effects of linguistic discourse from availability of the referent. *Journal of Memory and Language, 65,* 299-317. doi: 10.1016/j.jml.2011.04.010

Appendix A. Materials from Experiments 1.

Sample items from Experiment 1. After item 1, only items with a local disjunction are provided.

1. a. A guest talked to Bill or Sue, but I don't remember which (one).

   b. Bill or Sue talked to a guest, but I don't remember which (one).

2. A man argued with Peter or Janice, but I can't say which (one).

3. A preacher chatted with Bart or Connie, but I don't know which (one).

4. A lawyer phoned Susan or Paul, but I'm not sure which (one).

5. A professor challenged Lyn or Chuck, but I'm not certain which (one).

Appendix B. Materials from Experiment 2.

Sample materials from Experiment 2A-B. The experiment manipulated whether and how the remnant to the sluice was continued (Singular: *one* / Plural: *of them* / Indefinite: an Indefinite NP, like *guest*). Only conditions in which the disjunction appeared in the local position are provided. Spill over and final regions for Experiment 2A (sluicing) are presented before the slash (*because it was so long ago*), whereas those for Experiment 2B (sluicing clefts) are presented after the slash (*it was at this point*).

1. A guest talked to Bill or Sue, but I don't remember which {one | of them | guest} (because it was so long ago / it was at this point).
2. A man argued with Peter or Janice, but I can't say which {one | of them | man} (as I couldn't hear the conversation / it was in the end).
3. A preacher chatted with Bart or Connie, but I don't know which {one | of them | preacher} (since the church was so crowded/ it was anymore).
4. A lawyer phoned Susan or Paul, but I'm not sure which {one | of them | lawyer} (though we could probably find out / it was right now).
5. A professor challenged Lyn or Chuck, but I'm not certain which {one | of them | professor} (though I have a pretty good idea / it was at the moment).

Appendix C. Materials from Experiment 3.

Sample Biased (a) and Neutral (b) context sentences for the exit questionnaire in Experiment 3. Context sentences preceded the corresponding sluiced sentence from Experiment 2. The questions subjects saw in the exit questionnaire are provided below each item. The ratio of *Yes: No: Not Sure* responses is provided in parentheses.

1. a. The party was swarmed with people.

   Does the sentence imply that there were multiple guests at the party? (100:0:0)

   b. It was a particularly humid night.

   Does the sentence imply that there were multiple guests involved? (6:75:19)

2. a. There were lots of unfamiliar men at the convention.

   Does the sentence imply that there were several men at the convention? (100:0:0)

   b. The convention started early in the morning.

   Does the sentence imply that there were several men at the convention? (6:81:13)

3. a. The church was interviewing several members of the clergy.

   Is it likely that there were multiple preachers at the church? (69:6:25)

   b. The church was particularly beautiful that morning.

   Is it likely that there were multiple preachers at the church? (0:25:75)

4. a. The law firm was growing rapidly.

   Is it likely that there were multiple lawyers at the firm? (100:0:0)

   b. Apparently, there was a bad accident on the freeway.

   Is it likely that there were multiple lawyers involved? (19:50:31)

5. a. The conference was well attended by scholars all over the world.

Does the sentence imply that there are several professors involved? (69:19:13)

b. The article made a highly contentious claim.

Does the sentence imply that there are several professors involved? (0:81:19)