The Target of Relative Clause Extraposition:

An Experimental Investigation of C-command Effects

Pasha Koval[1*] and Jon Sprouse[1]

[1]Program in Psychology, Division of Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates.

*Corresponding author(s). E-mail(s): pasha.koval@nyu.edu;
Contributing authors: jon.sprouse@nyu.edu;

Abstract

This paper presents the results of a series of large-scale formal experiments that use two c-command diagnostics, Condition C and NPI licensing, to determine the structural position(s) of an extraposed relative clause (RC) with respect to its host and other constituents. The results indicate that RC extraposition alters the structural position of the RC relative to the constituents in the main clause but does not modify its position with respect to the host. This leads us to conclude that the constituent that undergoes RC extraposition is composed of both the RC and its host. In turn, this suggests that RC extraposition should be analyzed as an extraction rather than a subextraction phenomenon.

Keywords: syntax, extraposition, relative clause, Condition C, NPI licensing, reconstruction, connectivity, c-command, experimental syntax

1 Introduction

The main research question of this paper is what constituent undergoes relative clause (RC) extraposition. To de-trivialize it, consider an example of RC extraposition in (1).

(1) I interviewed [$_{\text{NP}}$ a guitarist] yesterday [$_{\text{RC}}$ that performed at Carnegie Hall].

Ross (1967) offers a surface-friendly account of (1): what moves to the right during RC extraposition is the RC itself.[1] Under this view, RC extraposition is a type of subextraction from DP. An alternative account is developed in Fox and Nissenbaum (1999, 2000) and Sportiche (2016, 2019). According to them, RC extraposition targets the entire DP comprising the host NP and the RC, but everything except the RC is left unpronounced in the derived position. The two approaches are illustrated in (2).

(2) a. [$_{\text{DP}}$ D NP ~~RC~~] .... RC *(subextraction account)*

b. [$_{\text{DP}}$ D NP ~~RC~~] .... [$_{\text{DP}}$ ~~D NP~~ RC] *(extraction account)*

This study uses two c-command diagnostics, Condition C and NPI licensing, to determine the structural position of an extraposed RC relative to parts of the same DP and to other constituents in the same clause. The subextraction account predicts that RC extraposition should have a uniform effect on the c-command relationships between the RC and both groups of constituents. In contrast, the extraction account expects an asymmetry: the RC and parts of the same DP move together and therefore are predicted to show the same c-command relationship as in the base position, while the c-command relationship between the RC and other constituents in the same clause should change as a result of RC extraposition. Our results corroborate the predictions of the extraction account. We observe that parts of the same DP c-command the RC regardless of RC extraposition, while the c-command relationship between the RC and other constituents from the same clause changes due to RC extraposition.

The rest of the paper is structured as follows: Section 2 provides the theoretical background. Sections 3 and 4 detail the Condition C and NPI licensing experiments,

[1] This paper only considers rightward movement theories of RC extraposition; see Koval and Sprouse (2023) for the justification.

respectively. Section 5 discusses the broader theoretical implications of the results. Section 6 concludes.

2 General background

The central assumption shared by both the subextraction and extraction accounts is that RC extraposition is an instance of rightward syntactic movement. Therefore, to compare the c-command predictions of both accounts, we must consider their possible interaction with three syntactic reconstruction/connectivity patterns: obligatory, optional, and anti-reconstruction. To this end, as we will see, the selection of the two c-command diagnostics is not arbitrary, as it enables us to differentiate between the three reconstruction patterns. An important caveat is that NPI licensing can also be subject to semantic reconstruction, which can further modify some of the predictions.

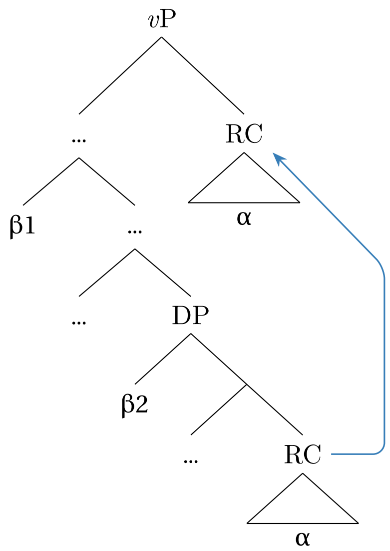
The discussion here focuses on RC extraposition from non-subject DPs, as they are not considered islands. In what follows, Condition C and NPI dependencies are compared across two different syntactic configurations. In the two configurations, the tail of the dependency (i.e. an R-expression for Condition C and a weak NPI for NPI licensing) occurs inside the RC, while the head (a coindexed pronoun and a polarity operator, respectively) occurs either inside the same DP or outside of it, but lower than the landing position of the extraposed RC (roughly, the edge of *v*P). We refer to the former configuration as *DP-internal* and to the latter as *DP-external*.

2.1 Extraction and subextraction theories of RC extraposition

Descriptively, rightward movement theories of RC extraposition can be classified based on two properties: the size of the target and the number of structural positions of an RC in a syntactic tree. Both can be determined by comparing c-command relationships of in-situ and extraposed RCs across DP-internal and DP-external configurations.

Consider the WYSIWYG theory of Ross (1967). According to it, RC extraposition moves the RC to the right. In (3), the RC is the target of RC extraposition and has two positions in a syntactic tree. Both DP-external and DP-internal dependency heads (shown as β1 and β2, respectively) c-command the dependency tail α inside the in-situ copy of the RC, while the extraposed copy appears outside of both c-command domains. Since the RC occupies two positions in a syntactic tree, Ross's theory is compatible with all three reconstruction patterns, but crucially it predicts the same effect of RC extraposition across both DP-internal and DP-external configurations.

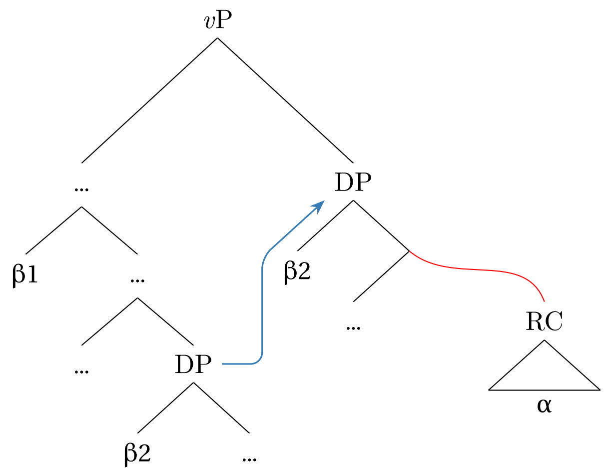
(3)



Turning to extraction theories, Fox and Nissenbaum (1999, 2000) argue that RC extraposition is generated as a sequence of two syntactic operations — the rightward movement of the DP via Quantifier Raising (QR) followed by the countercyclic Late Merge (LM; Lebeaux 1988) that merges the RC into the DP in its derived position at the right edge.[2] This places the RC in the derived position without creating a copy in

[2]To simplify the presentation, we first discuss a version of the QR+LM account that relies on the obligatory Late Merge (Stepanov, 2001b; Abe, 2018; Zyman, 2022).

the base position of the DP.[3] As a result, as shown in (4), the RC always appears in the c-command domain of β2 and is never found in the c-command domain of β1. This predicts an asymmetry between the two configurations: an anti-reconstruction pattern for the DP-external one and an obligatory reconstruction for the DP-internal one (but see Section 2.3 for another modification of this account using semantic reconstruction).

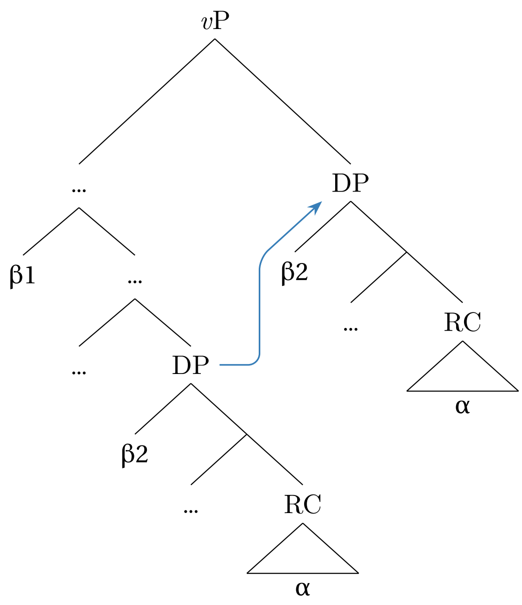
(4)


As discussed in Stepanov (2001b,a) and Sportiche (2019), the original definition of LM from Lebeaux (1988) makes LM optional for all adjuncts. Therefore, a QR+LM account of RC extraposition that adopts it allows RCs to be introduced in two ways: late merging the RC into its surface position (4), or merging the RC into the host DP in the base position and moving the resulting large DP to the right edge, via QR or Heavy NP Shift (5). According to this optional LM version of the QR+LM account, if at least one of the two derivations converges, RC extraposition is licensed. This also makes this account flexible with respect to the targets of RC extraposition (DP or DP+RC) and the number of positions of the RC (1 or 2). An important limitation of this account is that in both derivations the RC never appears outside of the DP, which

[3]See also Fox (2002, 2017a,b); Hulsey and Sauerland (2006); Takahashi and Hulsey (2009); Fox and Johnson (2016) for further discussion of different facets of this approach.

predicts the obligatory reconstruction pattern for the DP-internal configuration, while the DP-external one is compatible with any of the three reconstruction patterns.[4]

(5)



Lastly, Sportiche (2016) proposes to derive RC extraposition as an instance of Heavy NP Shift (HNPS).[5] During RC extraposition, the large DP that contains the RC moves to the right as shown in (5). According to this approach, the differences between HNPS and RC extraposition are created by the selective blindness principle called Neglect (6), which applies at both syntax-phonology and syntax-semantics interfaces, distinguishing possible and impossible PFs and LFs. (7) shows the only possible PF.

(6) Neglect

Any material at any interface can be ignored up to crash.

[4]Note that the obligatory and optional LM versions are both logically possible, since Fox and Nissenbaum (1999, 2000) do not specify which version of LM they adopt. In contrast, Fox and Johnson (2016) offer a multidominance equivalent of the obligatory LM version of the QR+LM account, since the extraposed RC always occupies a single position at the right edge, while its host NP1 has two mothers (see their pp. 6–7). Allowing NP2 to have two mothers instead would turn this into an equivalent of the optional LM version.

[5]See also Sportiche (2017, 2019). A very similar account is entertained in Wilder (1996) (his "R-account") but is rejected for conceptual reasons, since the rightward movement is not compatible with LCA.

(7) Possible PFs

a. [DP ~~RC~~] ...[~~DP~~ RC]

In turn, the choice of possible LFs for RC extraposition is guided by two principles in (8) and (9) and a structural requirement according to which the RC should always appear as an adjunct inside the DP at LF. The principle in (8), from Chomsky (1993), ensures that at least one copy of the RC appears at LF,[6] while the one in (9) says that the lower copy of DP should never be Neglect-ed at LF. (10) lists all possible LFs.

(8) Principle of Full Interpretation

Interpret every syntactic object.

(9) Local Predicate Saturation

At LF, a predicate must be locally saturated by its arguments.

(10) Possible LFs

a. [DP RC] ...[DP RC]

b. [DP RC] ...[~~DP~~ ~~RC~~]

c. [DP ~~RC~~] ...[DP RC]

d. [DP RC] ...[DP ~~RC~~]

As follows from (10), Sportiche's theory predicts the obligatory reconstruction pattern for the DP-internal configuration (since the RC never appears outside of the DP), while the DP-external configuration can show any of the reconstruction patterns.

Table 1 summarizes the resulting classification of the rightward movement theories of RC extraposition. As we can see, the two properties we started with distinguish theories that make different predictions: the obligatory LM version of the QR+LM theory

[6] An important distinction between the domains of (6) and (8) is that the former applies to individual links in a movement chain and the latter to entire chains, thus Neglect-ing one link in a movement chain at PF or LF does not violate (8) as long as there is another link of the same movement chain that is interpreted at that interface.

posits a single position for the RC and thus necessarily predicts the anti-reconstruction pattern for the DP-external configuration, while all two-position theories are compatible with all three reconstruction patterns for the same configuration. In turn, Ross's theory is the only one that moves the RC on its own and therefore predicts the uniform effect of RC extraposition across both DP-internal and DP-external configurations, while all other theories that target the DP+RC complex necessarily predict that the DP-internal configuration is restricted to the obligatory reconstruction pattern.

Table 1: Classification of rightward movement theories of RC extraposition according to the number of structural positions for the RC and the size of the moving constituent

	# of pos.	What moves
RC moves (Ross, 1967)	2	RC
QR + LM (Fox and Nissenbaum, 1999, 2000)		
+ obligatory LM (Stepanov, 2001b; Abe, 2018; Zyman, 2022)	1	DP
+ optional LM (Lebeaux, 1988)	1 or 2	DP or DP+RC
HNPS + Neglect (Sportiche, 2016, 2019)	2	DP+RC

The next section examines the use of Condition C and NPI licensing as c-command diagnostics and shows how their results can identify different reconstruction patterns.

2.2 The logic of the c-command diagnostics

There are two fundamental assumptions that underlie the use of Condition C and NPI licensing as c-command diagnostics. First, both Condition C and NPI licensing require the head of the dependency (a violating coindexed pronoun or a downward-entailing operator) to c-command the tail (a coindexed R-expression or a weak NPI); see Reinhart (1976, 1981) for Condition C and Klima (1964b); Ladusaw (1979); Barss and Lasnik (1986) for NPI licensing. In this way, any differences observed between extraposed and non-extraposed RCs with respect to both dependencies are understood as indicative of differences in c-command.[7] Second, Condition C and NPI licensing

[7] It is important to note that NPI licensing is usually analyzed in terms of scope, not c-command. An NPI must appear in the scope of a downward-entailing operator, which is only partially determined by syntax; see Chapter 3 of Ladusaw (1979) and Barker (2012) for a more detailed discussion. Here we abstract away

of weak NPIs do not exhibit a clause-mate restriction between the head and the tail, unlike, for example, NPI licensing of strong NPIs or Binding Theory Conditions A and B; see Reinhart (1976); Chomsky (1981) for Condition C and Ladusaw (1979); Progovac (1988, 1993); Hoeksema (2017) for NPI licensing. This assumption is crucial for testing RC extraposition since the head of the dependency can occur in the main clause while the tail always needs to stay inside the RC.

Building on the two assumptions above, the structural requirements for Condition C and NPI licensing are as follows:

(11) Condition C

An R-expression must appear outside the c-command domain of a coindexed pronoun.

(12) NPI licensing

An NPI must appear inside the c-command domain of a polarity operator.

Next, these requirements can be extended to movement chains where, as a result of movement, an R-expression or an NPI occupies multiple positions:[8,9]

---

from cases where the c-command requirement of NPI licensing may be construed as interacting with feature percolation. Consider the following contrast between the quantifiers *every* and *no* also due to Ladusaw (1979, p. 116). The former can only license an NPI *ever* in its scope (the RC), but not in the scope of its QP, while the latter can do both.

(i) a. Every student who had ever read anything on phrenology attended the lectures.
 b. *Every student who had attended the lectures had ever read anything on phrenology.

(ii) a. No student who had ever read anything on phrenology attended the lectures.
 b. No student who had attended the lectures had ever read anything on phrenology.

[8]The discussion here is couched in terms of the copy theory of movement (Corver and Nunes, 2007). Other movement theories (e.g. multidominance, Johnson 2020) can be plugged in instead, as long as they have a chain interpretation mechanism in place that ensures that a grammatical operation affects all links of a chain. So, for example, when an NPI is licensed in one position, all its copies are licensed as well.

[9]The details of the licensing environments for NPI and Condition C experiments can be found in their respective sections.

(13) Condition C in a movement chain

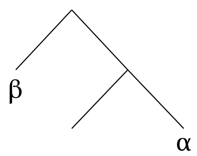
A copy of an R-expression must appear outside the c-command domain of a coindexed pronoun.

(14) NPI licensing in a movement chain

A copy of an NPI must appear inside the c-command domain of a polarity operator.

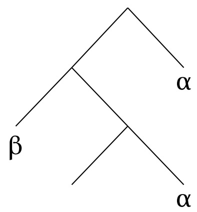
By combining (13) and (14), we can distinguish three scenarios depending on whether all, some, or none of the copies of a moved constituent α appear in the c-command domain of a licensor β. These scenarios are shown in (15) along with their standard reconstruction labels. If an R-expression inside α that is coindexed with a pronoun β triggers a Condition C violation, while an NPI inside α is licensed by a polarity operator in β, then all copies of α are in the c-command domain of β.In the case where there is no Condition C violation and an NPI is licensed, only some copies of α are in the c-command domain of β. Finally, if Condition C is not violated and an NPI inside α cannot be licensed by β, no copies of α are c-commanded by β.

(15) a. Obligatory reconstruction



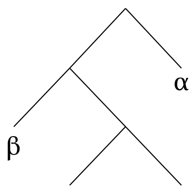
Condition C:	violation
NPI licensing:	licensed

b. Optional reconstruction



Condition C:	no violation
NPI licensing:	licensed

c. Anti-reconstruction



Condition C:	no violation
NPI licensing:	unlicensed

2.3 A caveat: semantic reconstruction for NPI licensing

Condition C is generally accepted to be evaluated at LF (Freidin, 1986; Lebeaux, 1988, 2000, 2009; Chomsky, 1993). Therefore, it can be used to identify the position of the moved element at LF and, through that, its reconstruction pattern (Heycock, 1995; Romero, 1998; Fox, 1999). The timing of NPI licensing is more controversial. Linebarger (1980); Mahajan (1990) and Uribe-Etxebarria (1995) provide a number of arguments that NPI licensing should be postponed at least until LF. Further, Moulton (2013) shows the pair in (16) in which the obligatory reconstruction would lead to a Condition C violation and the anti-reconstruction to an NPI licensing violation, yet he reports no disjoint-reference contrast in this pair. There are two possible ways to explain this: either CP fronting shows optional reconstruction or NPI licensing is not evaluated simultaneously with Condition C and should be moved even further to post-LF. The latter option makes NPI licensing subject to semantic reconstruction.

(16) a. That $John_i$ would ever lose a race, he_1 never expected.
b. That $John_1$ would lose a race, he_1 never expected.

The basic mechanism of semantic reconstruction is straightforward: the moved element leaves a variable in its base position which is bound by λ-operator inserted immediately below the derived position. An illustration is shown in (17). Semantic reconstruction is compatible with both extraction and subextraction accounts of RC extraposition. Crucially, if the variable appears in the scope of a downward-entailing operator, then the function that abstracts over that variable will be a downward-entailing function. Therefore, if NPI licensing is evaluated at this step, an NPI will be licensed regardless of the position of the RC at LF.

(17) ⟦Tim ordered every book today that Kate had ever recommended⟧^{g} =
$[\lambda f_{\langle e,t\rangle}.\forall x.$Tim ordered x s.t. x is a book and $f(x) = 1]$
(λy_e. Kate had ever recommended y)

This affects the predictions of the anti-reconstruction pattern that normally expects the absence of NPI licensing, as in (15c). Paired with semantic reconstruction, its predictions become indistinguishable from the optional reconstruction in (15b). Interestingly, this combo generates a new set of predictions even for the obligatory LM version of the QR+LM theory, which only puts the extraposed RC into the derived position. As a result of semantic reconstruction, both the DP and the late-merged RC should appear in the base position of the DP for the purposes of NPI licensing.[10]

2.4 Individual predictions of different theories

Table 2 shows all the possible combinations of a rightward movement theory and a reconstruction pattern, as well as the predictions that each combination makes for Condition C and NPI licensing in the DP-internal and DP-external configurations. Note that for Fox and Nissenbaum's and Sportiche's theories, the change in the reconstruction pattern is only visible in the DP-external configuration, as the DP-internal one is expected to only display the obligatory reconstruction. Semantic reconstruction is added as a suboption to theories where it would result in a different prediction.

[10]See Fox (2002) for a different perspective, in which the late-merged constituents are ignored during semantic reconstruction. The claim is based on the purported argument/adjunct reconstruction asymmetry for RC extraposition, which requires further exploration. If Fox is correct, it would rule out the semantic reconstruction modification of the obligatory LM account and, since the modified version generates the correct predictions, it would eliminate that account altogether. We have nothing more to say about this.

Table 2: The predictions of RC extraposition theories and reconstruction patterns for Condition C and NPI licensing across DP-internal and DP-external configurations

Theory	Condition C		NPI licensing	
	DP-internal	DP-external	DP-internal	DP-external
Ross (1967)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	no violation	no violation	licensed	licensed
+ anti-reconstruction	no violation	no violation	unlicensed	unlicensed
+ semantic reconstruction	no violation	no violation	licensed	licensed
Fox and Nissenbaum (1999, 2000) + Obligatory LM (Stepanov, 2001b; Abe, 2018; Zyman, 2022)				
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed
Fox and Nissenbaum (1999, 2000) + Optional LM (Lebeaux, 1988)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	violation	no violation	licensed	licensed
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed
Sportiche (2016, 2019)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	violation	no violation	licensed	licensed
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed

3 Condition C experiments

Experiments 1 and 2 test whether RC extraposition can avoid a Condition C violation in a DP-internal and a DP-external configuration, respectively. In Experiment 1, the binding dependency consists of a possessor of the host DP and a coreferential DP inside the RC. In Experiment 2, the head of the binding dependency is a theme argument of a ditransitive and the tail is a coreferential DP inside the RC adjoined to the goal argument. (18a) and (18b) show examples of target items from Experiments 1 and 2, respectively. The coreferential DPs are underlined.

(18) a. Amanda contacted his cousin ~~RC~~ yesterday [$_{\text{RC}}$ that babysat David].

b. Emily took him to a beach ~~RC~~ today [$_{\text{RC}}$ that Eric had never been to].

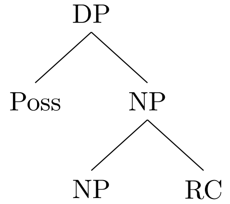
3.1 Background

3.1.1 The c-command relationship between a binder and a bindee

As argued in Reinhart (1976), for Condition C to apply, the binder must c-command the bindee. Since Experiments 1 and 2 use the Condition C violation to test RC extraposition, it is necessary to ensure that the binder c-commands the in-situ RC and, therefore, the bindee. At the same time, the binder should not appear too high in the structure, so as to c-command the extraposed RC.

Experiment 1 relies on two assumptions about the relative structural positions of possessors and in-situ RCs inside a DP in English. First, the possessor always appears at the DP level, as proposed by Abney (1987). Second, the in-situ position of a restrictive RC is adjoined to NP below D (Partee, 1975, p. 231).[11] Therefore, a possessor located within the DP layer c-commands an in-situ restrictive RC adjoined to the NP.[12] This relationship is shown schematically in (19).

(19)



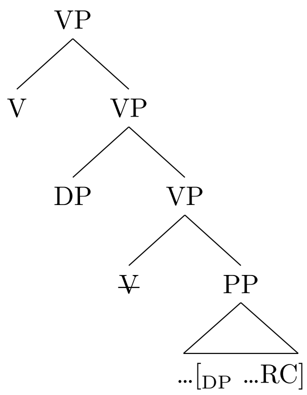
Experiment 2 assumes that there is a structural asymmetry between non-subject arguments in the oblique locative structure with ditransitive verbs. After Larson (1988), we assume the following structure for the ditransitive VP where the dative theme DP asymmetrically c-commands the goal PP:[13]

[11]See Stockwell et al. (1973) for a catalog of differences between restrictive and non-restrictive RCs in English.

[12]For example, we expect the following coindexing to be impossible:

(i) *Her$_i$ assistant that accompanied Teresa$_i$ smiled reassuringly.

[13]The base position of the verb inside VP is not relevant to us here; for a recent overview, see Harley and Miyagawa (2018).

(20) 

The following contrasts from Larson (1988) illustrate the asymmetry. We also conducted a follow-up experiment to test whether the asymmetry is present in the verbs used in Experiment 2. The details of this experiment can be found in Section 3.4.

(21) *Anaphor binding*

a. I presented/showed Mary$_i$ to herself$_i$.

b. *I presented/showed herself$_i$ to Mary$_i$.

(22) *Weak crossover*

a. Which check$_i$ did you send to its$_i$ owner?

b. *Which worker$_i$ did you send his$_i$ check to?

(23) *Superiority*

a. Which check did you send to who?

b. *Whom did you send which check to?

3.1.2 Using the coreference judgment task

An important methodological issue in experimental studies of binding and coreference is the choice of the dependent variable encoded in the experimental task. Theoretical syntactic literature systematically uses a procedure equivalent to the acceptability judgment task (Chomsky, 1981; Lasnik, 1989).[14] In this task, participants are asked to assess the acceptability of a sentence under a coreferential interpretation. In contrast, the experimental literature often employs the coreference judgment task (Gordon and Hendrick, 1997; Kazanina et al., 2007). In this task, participants are invited to assess the availability of a coreferential interpretation of a sentence. Experiments 1 and 2 both use the coreference judgment task. The instructions used in both are shown in (24). The purpose of this section is to clarify the reasons for choosing this task.

(24) **The coreference judgment task**

Your task is to determine whether the two underlined words could refer to the same person or whether they must refer to different people. You will rate this from -3 (they must refer to different people) to 3 (they could refer to the same person).

First, consider the sentence in (25). The coreferential reading is excluded because it violates Condition C, whereas the non-coreferential reading is unaffected and available.

(25) He paid for Timothy.

a. *He$_i$ paid for Timothy$_i$. *(coreferential reading)*

b. He$_j$ paid for Timothy$_i$. *(non-coreferential reading)*

In the acceptability judgment task, participants are required to focus only on (25a) and rate its acceptability using a scale (binary, n-point, etc.). It is crucial that (25b) is not part of that scale. The ability to focus on one interpretation while blocking out the other is the metalinguistic part of this task. The skill necessary to perform this part is usually taught (explicitly or implicitly) in introductory linguistics classes, often accompanied by the idea that the acceptability scores of different readings are

[14] At the same time, experimental binding studies using the acceptability judgment task are rare. One example is Temme and Verhoeven (2017).

independent. Thus, the status of the non-coreferential reading in (25b) does not change the rating of the coreferential reading in (25a) and vice versa. It seems premature to assume that all participants possess this skill. Thus, an experiment using this task should ensure that participants demonstrate some degree of metalinguistic awareness.

Second, during the course of a typical 2×2 experiment, the metalinguistic task must be repeated at least 35–40 times. Meanwhile, a non-coreferential interpretation is available in all sentences with a Condition C violation. This creates the risk that a participant, whenever feeling tired or distracted at any point during the experiment, may stop focusing on the coreferential reading and report the acceptability of a sentence *under any interpretation*, resulting in a false positive.

In contrast, in the coreference judgment task, the two readings appear on the same scale as shown in (24).[15] This eliminates the need for participants to differentiate between the two readings and isolate one of them, removing the metalinguistic component and simplifying the task.[16]

3.2 Methods

3.2.1 Experimental design

Both Condition C experiments use a full 2 × 2 factorial design. The two factors in this design are LINEAR ORDER and DEPENDENCY. The former manipulates the order of parts of a binding dependency between a Condition C-compliant “r-first” dependency where the head is an R-expression and the tail is a pronoun, and an inverse “p-first” dependency where the two are swapped, which should produce a Condition C

[15]The experimental literature knows some attempted variations of this. For instance, Stockwell et al. (2021) gave participants two distinct scales to evaluate the “naturalness” of each reading independently, while Keller and Asudeh (2001) offered a single scale for the non-coreferential reading only. It is clear that both reintroduce the metalinguistic aspect into the task.

[16]The coreference judgment task also has its limitations. For example, it cannot be used with most sentences testing Condition A, as the non-coreferential reading is ungrammatical, which creates a logical inconsistency for the used scale. To address this issue, Gordon and Hendrick (1997) suggested including an additional checkbox for participants to mark a sentence as ungrammatical, although this would make the task more complex. Therefore, in this paper, we decided against using any Condition A sentences.

violation. The latter controls the presence of RC extraposition. The abstract schemas for the four experimental conditions are shown in (26).

(26) a. [R-exp ... [$_{\text{RC}}$ ... pron ...]] r-first | in-situ

b. [R-exp ... ~~RC~~ ...] ... [$_{\text{RC}}$ ... pron...] r-first | ex-situ

c. [pron ... [$_{\text{RC}}$ ... R-exp ...]] p-first | in-situ

d. [pron ... ~~RC~~ ...] ... [$_{\text{RC}}$ ... R-exp ...] p-first | ex-situ

The central advantage of this design is that it separates two main effects: the cost of having a legal or an illegal binding dependency and the cost of an RC extraposition dependency. In this design, any improvement of the Condition-C-non-compliant binding dependency triggered by RC extraposition shows up as a sub-additive interaction. Figure 1 shows the mock plots with the two outcomes predicted by this design. Both plots show the main effects of LINEAR ORDER and DEPENDENCY represented by horizontal and vertical shifts between pairs of conditions. In the left panel, there is no interaction term, and the illegal binding dependency that violates Condition C is not ameliorated by RC extraposition. In contrast, the plot in the right panel shows a large subadditive interaction that indicates that in the p-first/ex-situ condition, RC extraposition puts the R-expression into a position outside of the c-command domain of the pronoun, creating a leftward-facing alligator mouth shape.

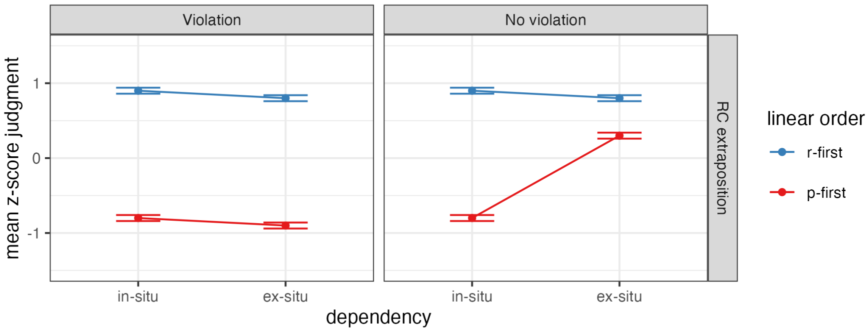


Fig. 1: Condition C (mock plots)

3.2.2 Materials

Experiments 1 and 2 investigate Condition C during RC extraposition. Experiment 1 studies RC extraposition from the direct object and Experiment 2 RC extraposition from the indirect object (goal). Both experiments employ a binding (c-command) dependency that includes a pronoun and an R-expression. A Condition C violation is created by making the pronoun the head of the binding dependency. In the pair of control conditions, the R-expression is the head of that dependency. The sets of sample conditions for the two Condition C experiments are shown in (27) and (28).

Experiment 1 uses the possessor as the head of the binding dependency to test Condition C during RC extraposition in a DP-internal configuration. A temporal adverb marks the right edge of the matrix clause in the ex-situ conditions. In the in-situ conditions, the same adverb appears at the left edge of the matrix clause to exclude its misattribution to the in-situ RC. Past perfect is used in the RC for sequence-of-tense reasons. The gender feature on pronouns and R-expressions is used to prevent other anaphoric elements from being included in the binding dependency. The head noun for the host DP is chosen to be animate and not (prototypically) unique (e.g. *aunt*, but not *father*) in order to allow for two restrictors (the possessor and the RC).

(27) Experiment 1: Condition C (*possessor*; DP-internal) × RC extraposition

a.	Yesterday Amanda contacted David's cousin [$_{\text{RC}}$ that babysat him].	r-first	in-situ
b.	Amanda contacted David's cousin ~~RC~~ yesterday [$_{\text{RC}}$ that babysat him].	r-first	ex-situ
c.	Yesterday Amanda contacted his cousin [$_{\text{RC}}$ that babysat David].	p-first	in-situ
d.	Amanda contacted his cousin ~~RC~~ yesterday [$_{\text{RC}}$ that babysat David].	p-first	ex-situ

Experiment 2 tests Condition C during RC extraposition in the DP-external configuration. The head of the binding dependency is the theme argument of a ditransitive, while the RC appears inside the goal argument. Similarly to Experiment 1, a temporal adverb marks the right edge of the matrix clause during RC extraposition. The same adverb appears at the left edge of the matrix clause in the in-situ conditions. The head of the host DP is indefinite to improve the plausibility of a sentence without context. The gender feature value of the parts of a binding dependency is different from all other anaphoric elements in a sentence to avoid misattribution.

(28) Experiment 2: Condition C (*ditransitive*; DP-external) × RC extraposition

a.	Today Emily took Eric to a beach [$_{\text{RC}}$ that he had never been to].	r-first	in-situ
b.	Emily took Eric to a beach today [$_{\text{RC}}$ that he had never been to].	r-first	ex-situ
c.	Today Emily took him to a beach [$_{\text{RC}}$ that Eric had never been to].	p-first	in-situ
d.	Emily took him to a beach today [$_{\text{RC}}$ that Eric had never been to].	p-first	ex-situ

3.2.3 Fillers and sanity check items

Both experiments include 10 filler items that encompass a wide range of availability of coreferential interpretation across different syntactic structures. Their primary goal is to distract participants from the experimental manipulation while encouraging them to use the entire scale. Additionally, filler responses are used for outlier detection purposes to identify sources of bias and error among participants.

Both experiments also include 3 pairs of sanity check items that test Condition B, Condition C, and Condition C under reconstruction. The checks are designed to test the experimental design and assess the accuracy of the collected responses. Each check focuses on a specific aspect of the data. Moreover, the checks show the relative acceptability levels of their respective violations, allowing us to distinguish no-violation cases from those where the ratings indicate a violation.

The Condition B sanity check shown in (29) tests whether participants pay attention to the structural positions of the head and tail of the binding dependency (i.e. attend to c-command) and also provides a reference point for a binding violation independent of the one tested in the experiment (Condition C). The Condition C sanity check in (30) confirms that participants attend to the task and pay attention to the categorical status and the structural and linear positions of anaphoric elements within a binding dependency. Finally, the sanity check that tests Condition C during the obligatory reconstruction in (31) serves two purposes. It provides a benchmark for the combined costs of reconstruction and a Condition C violation to ensure that there is enough room on the scale for both. It also shows the relative acceptability levels of both, allowing us to distinguish between a situation in which the rating suggests a violation and one in which there is no violation.

(29) Sanity check: Condition B

a. Hannah's aunt sometimes surprised her. — no violation

b. Hannah's aunt sometimes surprised her. — violation

(30) Sanity check: Condition C

a. Allison added that she liked reggae. — no violation

b. She added that Allison liked reggae. — violation

(31) Sanity check: Reconstruction + Condition C

a. Bella said that [$_{PP}$ in front of him], the tailor saw a moth ~~PP~~. — no violation

b. Bella said that [$_{PP}$ in front of the tailor], he saw a moth ~~PP~~. — violation

3.2.4 Survey composition

Each binding experiment consisted of 8 sets of lexically matched experimental items. Participants were shown 2 items per condition. The survey included a total of 33 items, which were presented in the following order: 9 practice items in a fixed order, followed by 8 experimental items, 10 filler items, and 3 sets of sanity check items (2 items each) presented in a pseudorandomized order. Each participant saw a comparable number of Condition B and Condition C sentences. Both experimental and sanity check items were distributed among the experimental lists using a Latin square procedure. To control for order effects, 8 lists of each experiment were shown in 4 counterbalanced orders.

3.2.5 Participants and presentation

A total of 240 participants were recruited for both experiments, with 120 participants assigned to each of them. This sample is projected to yield 100% statistical power for a 7-point scale and medium-size effects (Sprouse and Almeida, 2017). The participants were instructed to evaluate whether the two underlined DPs can refer to the same person (3) or if they must be different people (-3). Each sentence was presented on a separate screen and had a separate scale next to it.

The experiments were hosted online on a survey platform Qualitrcs. The recruitment was carried out through a crowd-sourcing platform Amazon Mechanical Turk and a recruitment facilitation service CloudResearch. All participants were self-reported native English speakers and were compensated for their time at an hourly rate of $15 per hour with an estimated completion time of 6 minutes. Each participant saw only one list of one experiment and all the experimental conditions in it.

3.2.6 Analysis

The results were standardized to z-scores to remove scale bias. Three methods were employed to detect and exclude uncooperative participants: Tukey's inner fences (Tukey, 1977), the sum of squared errors (SSE), and Iglewicz and Hoaglin's exact fit test (Iglewicz and Hoaglin, 1993). This led to the removal of 3 participants from Experiment 1 and 6 participants from Experiment 2. We then calculated the empirical floor and ceiling following Al-Aqarbeh and Sprouse (2022); Fukuda et al. (2022), which would help to identify overpowering main effects obscuring the interaction terms.

The analyses were performed using R version 4.1.1 (R Core Team, 2021). Linear mixed-effects models were constructed for each experiment using the `lme4` package (Bates et al., 2015) with LINEARD ORDER and DEPENDENCY as fixed effects and PARTICIPANT and ITEM as random effects. p-values were computed using the `lmerTest` package (Kuznetsova et al., 2017), and Bayes factors of the BF10 type were derived using the `BayesFactor` package (Morey and Rouder, 2018). For ease of exposition, the interaction plots include both the interaction term p-value and the BF10 value.

BF_{10} shows the ratio between the likelihood of the data under the experimental (H1) and null (H0) hypotheses, allowing for the evaluation of H1 and H0 directly. For instance, $BF_{10} = 3$ suggests that the data is 3 times more likely under a theory that predicts the interaction (H1) than under one with no interaction term (H0). If we find

BF_{10} = 1/3, we conclude that the data is 3 times more likely under H0 than under H1. This helps to distinguish between "null" results that support H0 and those that are fundamentally inconclusive. The conventional thresholds for *p*-values and BF_{10} are adopted from Neyman and Pearson (1928) and Jeffreys (1939), respectively. Using *p*-values and BF_{10} values together allows us to identify the following three patterns:

1. A *p*-value $< .05$ and $BF_{10} > 3$ indicate that RC extraposition ameliorates a Condition C violation.
2. A *p*-value $> .05$ and $BF_{10} < 0.33$ indicate that RC extraposition does not affect a Condition C violation.
3. A *p*-value $> .05$ and $0.33 < BF_{10} < 3$ signal the lack of strong support for either hypothesis.

3.3 Results

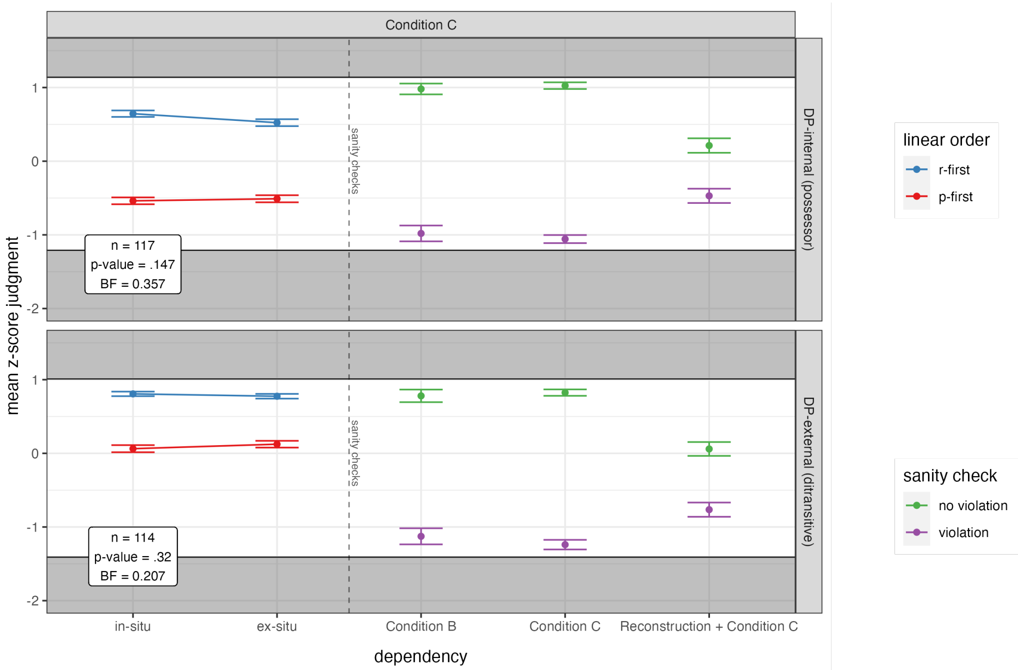


Fig. 2: Interaction plots for both Condition C experiments

Figure 2 contains the results of both Condition C experiments. In the top panel, which shows the results of Experiment 1, we observe that there is no "inverse" alligator mouth shape within the experimental conditions. Both Condition B and Condition C sanity checks confirm that the participants paid attention to the structural and linear properties of both elements in a binding dependency ($p < .001$ for both). The Reconstruction + Condition C sanity check shows that the combined costs of reconstruction and a Condition C violation had enough space and would not cause the floor effect ($p < .001$). Importantly, the size of the main effect of LINEAR ORDER is greater than the cost of a Condition C violation during reconstruction in the sanity check. It is also greater than the cost of having a grammatically correct forward vs. backward binding dependency, as seen in the comparison of the Condition C/no violation and Reconstruction + Condition C/no violation conditions. This suggests that the main effect of the LINEAR ORDER factor in Experiment 1 is created by a Condition C violation. Paired with the lack of an "inverse" alligator mouth shape, it indicates that in a DP-internal configuration, RC extraposition does not bleed Condition C. Both sets of statistical tests agree that there is no effect of RC extraposition on the Condition C violation, since both find no evidence of the presence of a subadditive component feeding into the interaction term at significance levels $p < .05$ and $BF_{10} > 3$. The BF_{10} value strongly supports the null hypothesis that there is no interaction ($BF_{10} < 0.33$).

The bottom panel contains the results of Experiment 2. Here we also observe no "inverse" alligator mouth shape, but both pairs of conditions appear in the top half of the scale and are much closer to each other than in Experiment 1. The Condition B and Condition C sanity checks show that the participants were aware of the structural and linear positions of the parts of a binding dependency ($p < .001$ for both). Also, the Reconstruction + Condition C sanity check shows that there is enough space should Condition C occur ($p < .001$). However, the comparison of two sanity check conditions (Condition C/no violation and Reconstruction + Condition C/no violation) shows that the cost of having a grammatical forward vs. backward binding dependency is approximately equivalent to the main effect of the LINEAR ORDER factor (note that both occur at the same acceptability levels), suggesting that there is no Condition C violation in the violation pair of experimental items in Experiment 2. Both sets of statistical tests suggest that there are no factors contributing to the interaction term. In Experiment 2 we did not observe an acceptability level indicative of a Condition C violation in either of the p-first conditions. This could be explained in two ways.

The first explanation suggests that the syntactic structure used in Experiment 2 does not involve a binding violation in the first place and thus cannot serve as a test for its obviation by RC extraposition. This could be the case if the proposal from Larson (1988) about an asymmetric c-command relationship between the two non-subject arguments in the oblique locative construction is incorrect.[17]

According to the second explanation, both p-first conditions include an obviation of a Condition C violation. This is expected if RC extraposition can occur string-vacuously as shown in (32). If this explanation is on the right track, it suggests that RC extraposition can obviate a Condition C violation in a DP-external configuration.[18]

(32) Today Emily took+T [$_{VP}$ him to a beach ~~RC~~] [$_{RC}$ that Eric had never been to].

In summary, Experiment 1 suggests that RC extraposition has no effect on a Condition C violation in a DP-internal configuration. However, the results of Experiment 2 are ambiguous. It is possible that the used syntactic configuration is not a valid test for Condition C, or it could be that RC extraposition can obviate the Condition C violation in a DP-external configuration, and, surprisingly, RC extraposition can apply string-vacuously. In order to distinguish between the two alternative explanations for

[17] It is a separate question whether certain simplex Ps are "true" prepositions or case markers (and therefore should not "count" for c-command), see Bruening (2018) for discussion and references.

[18] To test this explanation directly, one could try to block RC extraposition from happening string-vacuously by introducing another constituent (e.g. a result clause or a Heavy NP) at the right edge of the main clause. If this results in a Condition C violation, then this explanation is valid, and RC extraposition can occur string-vacuously. We defer testing this prediction to future research.

Experiment 2, we conducted a follow-up experiment discussed in the next section that tests whether Condition C is violated in a configuration with two non-subject arguments of a ditransitive in the oblique locative construction.

3.4 A follow-up: Condition C violation in ditransitives

The main purpose of this follow-up is to test the structural assumption underlying Experiment 2. Namely, we want to determine whether there is an asymmetric c-command relationship between the two non-subject arguments of a ditransitive in the oblique locative construction. If there is no structural asymmetry, the lack of a Condition C violation in Experiment 2 would indicate that this configuration is not a valid test for RC extraposition. Conversely, if we find that the theme c-commands the goal in that class of ditransitives, we can conclude that the lack of a Condition C violation in Experiment 2 was due to RC extraposition happening string-vacuously.

Experiment 3 uses Condition C to test whether the theme argument c-commands the goal argument. Similarly to previous experiments, it uses a full 2 × 2 factorial design. The design includes two factors, LINEAR ORDER and C-COMMAND. The former controls the order of elements in a binding dependency. In the r-first order, the R-expression is the head of the dependency and the pronoun is its tail. In the p-first order, the two are swapped. The latter factor separates the binding pair in which one anaphoric elements c-commands the other from the accidental coreference pair in which neither coreferential element c-commands the other. Experiment 3 uses the same set of verbs as Experiment 2. (33) contains a sample set of experimental items.

(33) Experiment 3: Condition C in ditransitives (linear order × c-command)

a.	Today Emily took Eric's sister to his neighborhood park.	r-first	no c-command
b.	Today Emily took Eric to his neighborhood park.	r-first	c-command
c.	Today Emily took his sister to Eric's neighborhood park.	p-first	no c-command
d.	Today Emily took him to Eric's neighborhood park.	p-first	c-command

In this design, the structural signature of a Condition C violation (an inverse anaphoric dependency in the presence of c-command) is distributed between two factors. LINEAR ORDER controls the order of elements in the anaphoric dependency, while C-COMMAND distinguishes the binding pair (c-command) from the accidental coreference pair (no c-command). The Condition C violation should feed into the interaction term, thus creating an open alligator mouth shape facing right.

The experiment consists of 8 sets of experimental items. All other items in the survey (i.e. fillers, sanity checks, practice items, and anchor items) are the same as in Experiment 2, as are the survey organization and presentation. The task, recruitment setup, remuneration rate, and outlier detection and statistical analysis procedures are identical to Experiments 1 and 2. We recruited 120 self-reported native English speakers, of which 2 were flagged as outliers and removed.

The results are shown in Figure 3. We observe a large dispreference for inverse anaphoric dependencies in which an R-expression follows a coreferential pronoun across both accidental coreference (no c-command) and binding (c-command) pairs. Additionally, in the r-first order, we find a minor boost in acceptability between accidental coreference and binding conditions, which could be due to the preference for

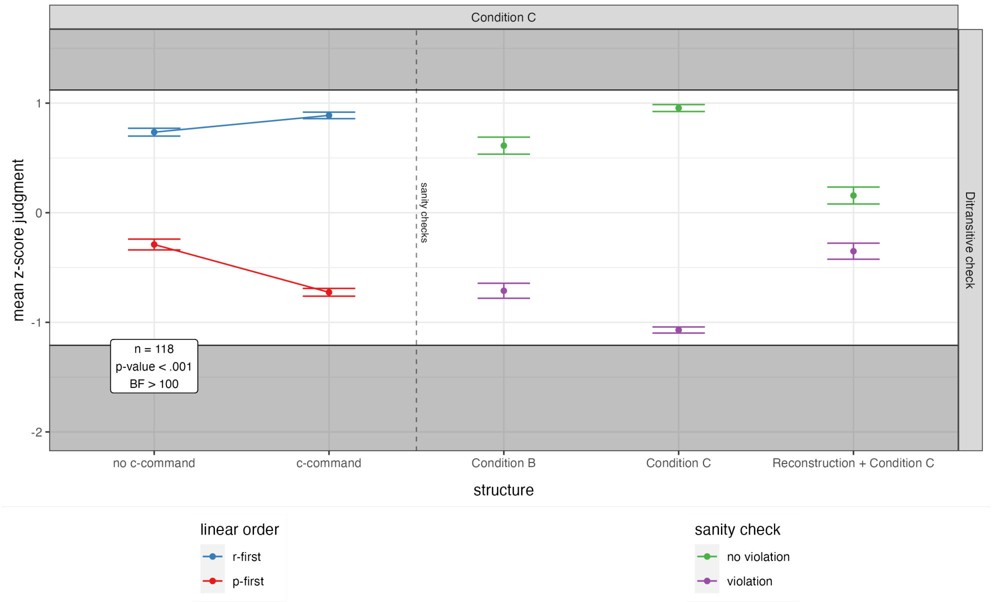


Fig. 3: Interaction plot for the Condition-C-in-ditransitives check

binding over coreference, or it could track the size of a DP in the structure. Finally, the binding condition in the p-first order shows an additional large penalty, which is indicative of a Condition C violation. The presence of a Condition C violation is corroborated by both sets of statistical analyses at the significance levels $p < .05$ and $BF_{10} > 3$. Thus, we can conclude that there is a structural asymmetry between non-subject arguments of a ditransitive in the oblique locative construction. This suggests that the same structural configuration in Experiment 2 is a valid test and therefore should trigger a Condition C violation, which is then obviated by RC extraposition. Therefore, the results of Experiment 2 should be interpreted as indicating that RC extraposition can apply string-vacuously to ameliorate a Condition C violation.

3.5 Discussion

The results of Condition C experiments are summarized in Table 3. We find that RC extraposition does not help the Condition C violation in the DP-internal configuration, but it does ameliorate the Condition C violation in the DP-external configuration. Moreover, we also now have evidence from Condition C that RC extraposition can apply string-vacuously.

Table 3: Summary of experimental results testing the effect of RC extraposition on Condition C in English

#	C-command diagnostic	Syntactic configuration	Result
1	Condition C	DP-internal	violation
2	Condition C	DP-external	no violation

4 NPI experiments

Experiments 4 and 5 test the acceptability of sentences with an NPI dependency that is interrupted (or not) by RC extraposition. Experiment 4 tests a DP-internal configuration that includes a universal quantifier *every* inside the host DP and a weak NPI *ever* inside the RC as, respectively, a head and tail of an NPI dependency. Experiment 5 uses the matrix negation marker *n't* and the same weak NPI *ever* in a DP-external configuration.

4.1 Background

4.1.1 The scope of polarity operators

The target sentences for both experiments are shown in (34). Both the universal quantifier *every* and the matrix negation *n't* are widely recognized as polarity operators that can license NPIs; see, for example, Klima (1964a); Ladusaw (1979); Linebarger (1980); von Fintel (1999) among many others. An important prerequisite for studying the reconstruction pattern of RC extraposition is that the extraposed RC appears outside the c-command domain of these operators.[19]

[19]For the two polarity operators used here, c-command is a good approximation of their scope, but see fn. 7 and references therein about the differences between scope and c-command.

(34) a. Emily answered [$_{\text{DP}}$ every question ~~RC~~] today [$_{\text{RC}}$ that Jon had ever asked in the group chat].

b. I [didn't respect [$_{\text{DP}}$ the players ~~RC~~]], for most of my life, [$_{\text{RC}}$ that had ever faked a serious injury].

A potential concern with using quantifiers to test NPI licensing is that in English they can undergo QR, which could potentially widen their scope far enough to include the extraposed RC. However, as shown in Ladusaw (1979), *every* only licenses NPIs in its restrictor argument (the sister of Q) but not in its non-nuclear scope (the sister of a QP formed by it). As a result, even if *every* moves at LF to a higher position via QR, it cannot license an NPI inside an extraposed RC from its new position.

In turn, the matrix negation marker *n't* c-commands the entire *v*P, which includes not only the base position of an extraposed RC, but also one of the (potential) landing sites of RC extraposition at the edge of *v*P (see Baltin 1981; Culicover and Rochemont 1990). To ensure that the extraposed RC leaves its base position and appears outside the scope of *n't* at the same time, Experiment 2 uses a second scope-taking element inside the *for*-phrase placed at the right edge of a clause that appears outside of the scope of *n't*. Thus, when an extraposed RC crosses it, it necessarily leaves the scope of *n't*. Consider the sentence in (34b). Its only interpretation is that for the majority of the speaker's (*I*) life, they have had no regard for a certain category of players. A temporal *for*-phrase with *most* appears outside of the scope of *n't* (*most* > *n't*).

We can verify that this interpretation of (34b) reflects the relative scope of *n't* and *most* and is not created from the inverse scope reading (*n't* > *most*) with the help of the pragmatic mechanism of negative strengthening (Horn, 1989). The inverse scope reading can be paraphrased as follows: it is not true that for the majority of the speaker's life, they deeply respected a certain category of players. During the negative strengthening, the lack of deep respect is reinterpreted ("strengthened") to mean the negative maximum, i.e. the lack of any respect. The resulting interpretation is as follows: it is true that for the majority of the speaker's life, they lacked any respect for a certain category of players. The key difference from the scope reading (*most* > *n't*) is that the implicature created by the negative strengthening is defeasible. For example, in (35) the implicature ("dislike cronuts") can be explicitly denied.

(35) I don't like cronuts, but I don't dislike them either.

However, trying to defuse the same implicature in (34b) produces a non-sequitur:

(36) I didn't respect the players, for most of my life, that had ever faked a serious injury, #but I didn't lack any respect for them either.

This suggests that the only reading of (34b) is not created by the negative strengthening and simply reflects the scope relations of the *for*-phrase and *n't*.

4.1.2 Using *ever* as a weak NPI

Overfelt (2015) conducted the only known acceptability judgment experiment that examined NPI licensing during RC extraposition. He tested a DP-internal configuration with *every* as the polarity operator, but instead of *ever*, he used another weak NPI, *any*. A sample target sentence from his experiment is shown in (37). His results indicate that RC extraposition does not interrupt NPI licensing in a DP-internal configuration.

(37) Park rangers removed every camper ~~RC~~ yesterday [$_{RC}$ who was at any of the sites with significant flooding].

As discussed in Gajewski (2016), a potential issue with using a weak NPI *any* in an experimental setting is that it also has a free-choice reading, which does not require a licensor and therefore must be excluded independently; see (Hoeksema, 2017) who shows that the free-choice reading of *any* is generally available in the restrictor argument of *every*. To avoid this confounding factor, Overfelt (2015) places most instances of *any* in his experimental items in partitive contexts following Dayal (2009), who argues that free choice items are impossible in partitive contexts.

In contrast, a weak NPI *ever* does not have a free-choice reading, thus eliminating the problem. Therefore, both experiments reported here use *ever* instead of *any*.

4.1.3 Using the existence presupposition as a disruptor for NPI licensing

Another potential issue to address is that in Experiment 2 the host of the RC is a definite plural DP, e.g. *the players* in (34b). Guerzoni and Sharvit (2007) argue that the plural definite description $\textit{the}_{\text{PL}}$ can license NPIs in its restrictor, while the singular $\textit{the}_{\text{SG}}$ cannot. If this is the case, using $\textit{the}_{\text{PL}}$ creates a confound, as it adds a new DP-internal NPI licensor to the structure.

Giannakidou (2002); Hoeksema (1986), and Homer (2010) challenge this claim, suggesting instead that the presence of the existence presupposition is what determines the possibility of NPI licensing by a definite description. Furthermore, Gajewski (2016) shows the experimental evidence suggesting that the grammatical number of a definite NP only has a moderate effect on NPI licensing, unlike prototypical NPI licensors. He concludes that such cases have to be separated from the core cases of NPI licensing and possibly handled by a different mechanism.

To mitigate the effect of $\textit{the}_{\text{PL}}$, we adopt a technique described in Homer (2010, Ch. 3, Appendix A). Homer argues that $\textit{the}_{\text{PL}}$ can license NPIs in its restrictor, but only in the absence of the existence presupposition. It follows that introducing an existence presupposition is going to disrupt that licensing relationship.[20] For this reason, Experiment 2 only uses matrix verbs that carry an existence presupposition for its direct object, e.g. *respect* in (34b).

[20] At the same time, importantly, disrupting the licensing relationship between $\textit{the}_{\text{PL}}$ and *ever* is not going to turn $\textit{the}_{\text{PL}}$ into an NPI anti-licensor since it is not left upward monotone.

4.2 Methods

4.2.1 Experimental design

Both NPI experiments use a full 2 × 2 factorial design. The factors are POLARITY ITEM and DEPENDENCY, each consisting of two levels. The factor POLARITY ITEM regulates the presence of a weak NPI *ever* inside the RC. DEPENDENCY manipulates the placement of the RC between an in-situ and an extraposed position. (38) shows the abstract schemas for the four experimental conditions.

(38)	a.	[*Op* ... [$_{\text{RC}}$ ...]]	no NPI	in-situ
	b.	[*Op* ... ~~RC~~ ...] ... [$_{\text{RC}}$ ...]	no NPI	ex-situ
	c.	[*Op* ... [$_{\text{RC}}$ ... NPI ...]]	NPI	in-situ
	d.	[*Op* ... ~~RC~~ ...] ... [$_{\text{RC}}$ ... NPI ...]	NPI	ex-situ

The primary advantage of this design lies in its ability to isolate the NPI licensing violation from the main effects of having an NPI dependency and an RC extraposition dependency in the structure. In this design, the cost of having an unlicensed NPI in the structure feeds the interaction term.[21]

Figure 4 shows two mock plots that illustrate two outcomes predicted by this design. The left panel demonstrates a scenario in which RC extraposition does not interact with NPI licensing. The main effects of POLARITY ITEM and DEPENDENCY appear as horizontal and vertical shifts between pairs of conditions, thereby preserving symmetry. Since RC extraposition does not interrupt the NPI dependency in the

[21] More carefully, the interaction term in length-based designs, such as the one shown in (38), has at least two possible sources. The first source is the cost of having an unlicensed NPI in the structure. The second is the cost of increasing the structural and linear lengths of the NPI dependency from the NPI/in-situ to the NPI/ex-situ condition. Although there are no theoretical claims suggesting that length is a factor that influences NPI licensing, this possibility aligns with our post-hoc hypothesis that some participants strongly prefer a structurally local NPI dependency; see Section 4.4. No interaction term was found in either of our NPI experiments, but if it were, a follow-up would be needed to differentiate between the two sources.

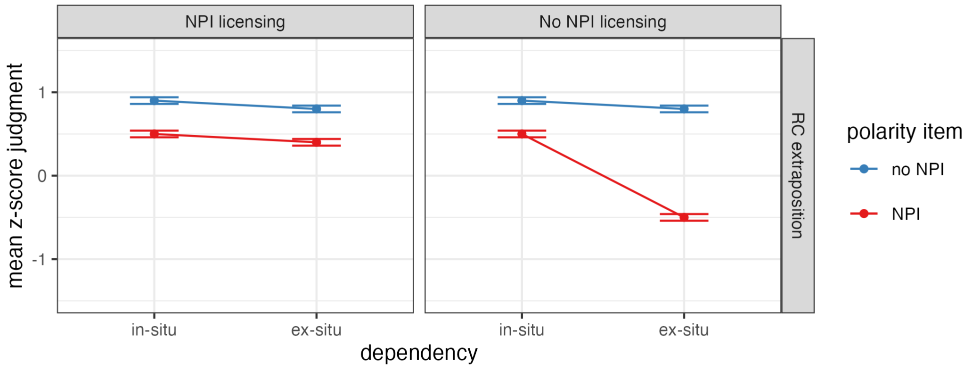


Fig. 4: NPI licensing (mock plots)

NPI/ex-situ condition, there is no interaction term that would break that symmetry. Moreover, the absence of an interaction term suggests that, under reconstruction, the length of an NPI dependency has no effect on acceptability. Otherwise, a longer NPI dependency in the NPI/ex-situ condition would be rated lower than a short NPI dependency in the NPI/in-situ condition. The plot in the right panel is expected if RC extraposition does interrupt NPI licensing. The characteristic left-facing alligator mouth shape, familiar from island experiments, is created by combining the two main effects with the interaction term, which suggests that the NPI/ex-situ condition contains an unlicensed NPI (and/or the NPI dependency in that condition is longer).

4.2.2 Materials

Experiments 4 and 5 examine NPI licensing during RC extraposition. Both experiments test RC extraposition from the direct object and use a weak NPI *ever* across all sets of experimental items. In both experiments, the pair of target conditions includes both ingredients of an NPI dependency, the polarity operator and the NPI, whereas the pair of grammatical controls only includes the polarity operator but not the NPI.

Experiment 4 uses the universal quantifier *every* as a polarity operator to test NPI licensing during RC extraposition in a DP-internal configuration. The right edge of the matrix clause in the ex-situ conditions is marked with a temporal adverb, whereas in the in-situ conditions, the same adverb is placed at the left edge of the matrix clause to avoid its misattribution to the in-situ RC. Past perfect is used in the RC because it sounds natural with *ever* and fits the tense and aspect of the main clause. A sample set of experimental items is shown in (39).

(39) Experiment 4: NPI licensing (*every*; DP-internal) × RC extraposition

a.	Today Emily answered every question [$_{\text{RC}}$ that Jon had asked in the group chat].	no NPI	in-situ
b.	Emily answered every question ~~RC~~ today [$_{\text{RC}}$ that Jon had asked in the group chat].	no NPI	ex-situ
c.	Today Emily answered every question [$_{\text{RC}}$ that Jon had ever asked in the group chat].	NPI	in-situ
d.	Emily answered every question ~~RC~~ today [$_{\text{RC}}$ that Jon had ever asked in the group chat].	NPI	ex-situ

Experiment 5 tests NPI licensing during RC extraposition in a DP-external configuration using the matrix negation marker *n't* as a polarity operator. The right edge of the matrix clause in ex-situ conditions is marked with a temporal *for*-phrase that outscopes the matrix negation. Therefore, when an RC extraposes across the *for*-phrase, it leaves the scope of *n't*. The same *for*-phrase appears on the left edge of the matrix clause in the in-situ conditions to avoid prosodic and semantic complications. All matrix verbs carry the existence presupposition for the direct object, thus disrupting NPI licensing by the definite plural *the*$_{\text{PL}}$ inside the host DP. Similarly to the previous experiment, the tense and aspect specifications of the RC are selected to sound natural with the NPI *ever* and to align with the tense and aspect values of the main clause. (40) shows a sample set of experimental items for this experiment.

(40) Experiment 5: NPI licensing (*n't*; DP-external) × RC extraposition

a. For most of my life, I didn't respect the players [$_{RC}$ that had faked a serious injury]. — no NPI | in-situ

b. I didn't respect the players ~~RC~~, for most of my life, [$_{RC}$ that had faked a serious injury]. — no NPI | ex-situ

c. For most of my life, I didn't respect the players [$_{RC}$ that had ever faked a serious injury]. — NPI | in-situ

d. I didn't respect the players ~~RC~~, for most of my life, [$_{RC}$ that had ever faked a serious injury]. — NPI | ex-situ

4.2.3 Sanity check items

To ensure that participants are attending to the task and not simply ignoring NPIs (especially unlicensed), both NPI experiments include sanity check items organized into minimal pairs. A sample minimal pair is shown in (41). The weak NPI *ever* is either licensed in the scope of a negative quantifier *no* or unlicensed in the scope of *some*. We created 8 lexically matched minimal pairs using the same weak NPI *ever* as the experimental items. Each experiment includes 4 sanity check items. Additionally, organizing sanity check items into minimal pairs allows us to measure the effect size of an NPI licensing violation that cannot be improved by movement/reconstruction.

(41) Sanity check items

a. Nobody has ever declined a Michelin star. — licensed

b. Somebody has ever declined a Michelin star. — unlicensed

4.2.4 Anchor items, practice items, and fillers

The NPI experiments include the same 3 anchor items, 9 practice items, and 14 fillers. All sentences and their expected ratings on the 1-7 scale are adopted from Sprouse et al. (2013). The anchor items are included in the instructions with their ratings, and participants do not rate them. The purpose of the anchor items is to demonstrate the use of the scale. The practice items are chosen to incorporate all 7 points of the scale, with the endpoints appearing twice. The fillers are used to distract participants from the experimental manipulation, while giving them the opportunity to use the entire scale. The responses to the fillers are also used to identify uncooperative participants.

4.2.5 Survey construction and presentation

The NPI experiments comprised 8 sets of lexically matched experimental items and 8 pairs of sanity check items, all of which were distributed among the experimental lists following a Latin square procedure. Each participant saw two tokens for each of the four experimental conditions and two tokens for each of the two types of sanity check items. The survey consisted of a total of 35 items, including 9 practice items presented in a fixed order and 8 experimental items, 4 sanity check items, and 14 fillers in a pseudorandomized order. To control for order effects, the experiment employed 4 lists presented in 4 counterbalanced orders. Each participant saw only one list per experiment and all the experimental conditions in it. Participants were instructed to judge each sentence on a scale from 1 (very bad) to 7 (very good). Each sentence appeared on a separate screen and had its own individual scale next to it.

4.2.6 Participant recruitment

A total of 360 participants were recruited for the NPI experiments. Initially, a sample of 240 individuals was recruited, with 120 assigned to each experiment. Furthermore, a second batch of 120 participants was recruited for Experiment 2 during post hoc analysis; see Section 4.4. According to Sprouse and Almeida (2017); Marty et al. (2020), the sample size of 120 participants is projected to yield close to 100% statistical power for the 7-point scale acceptability task for medium-size effects, such as NPI licensing during reconstruction. All participants were self-reported native English speakers and were compensated for their time at an hourly rate of $15 with an estimated completion time of 6 minutes. The experiments were hosted online on the survey platform Qualitrcs. Participants were recruited through the crowd-sourcing platform Amazon Mechanical Turk by means of a recruitment facilitation service CloudResearch.

4.2.7 Analysis

To control for scale bias, all results were normalized by converting them to z-scores prior to analysis. All outlier detection and statistical analysis procedures are identical to those described in Section 3.2.6. The outlier detection methods together flagged for removal 1 participant in Experiment 4 and 3 participants in each of the two batches of participants (i.e. to a total of 6 participants) in Experiment 5.

4.3 Results

The results of the NPI experiments are presented in Figure 5. In Experiment 4 shown in the left panel, we find no "alligator mouth" and therefore no sign of a superadditive interaction. The results of visual observation are corroborated by two sets of statistical tests. At the significance levels of $p < .05$ and $BF_{10} > 3$, both do not support the alternative hypothesis (that there is an interaction term). Furthermore, the Bayes Factor of $BF_{10} < 0.33$ indicates that there is strong evidence for the null hypothesis. In summary, our results strongly suggest that there is no interaction term and therefore NPI licensing by *every* is not interrupted by RC extraposition.

The right panel displays the results of Experiment 5. The shape we observe does not match either of the patterns from Figure 4 that are predicted by this design. Therefore, we cannot interpret these results, including the statistical tests. A more detailed analysis of this matter is deferred to Section 4.4.

The results of the sanity check suggest that in both experiments, the participants paid close attention to the licensing environments of NPI items (Exps. 4 & 5: $p < .001$).

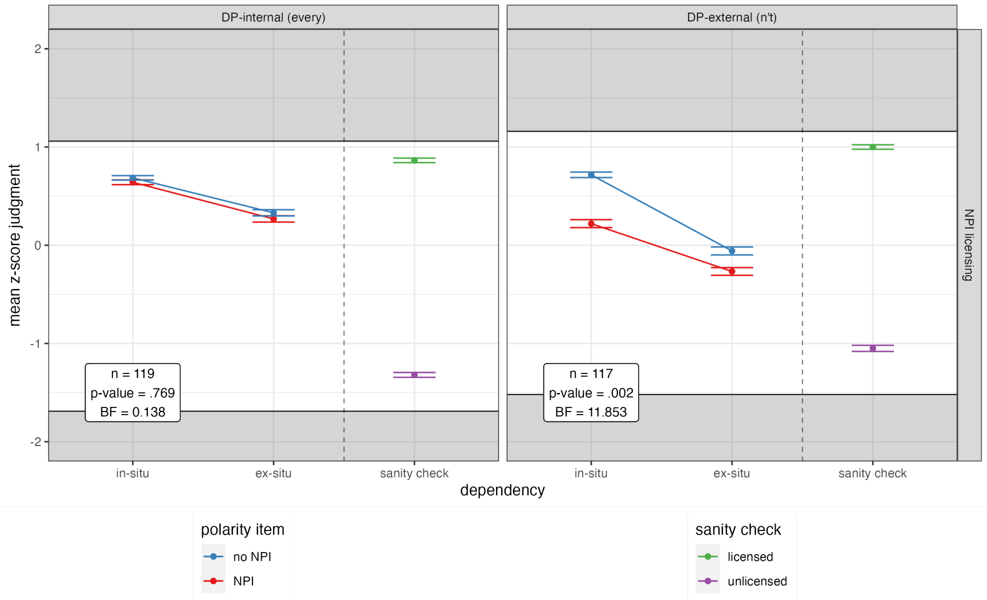


Fig. 5: Interaction plots for both NPI experiments (prior to clusterization)

4.4 The post-hoc cluster analysis of Experiment 5

In Experiment 5 in the right panel of Figure 5 the pair of NPI conditions shows an unexpected subadditive activity. On closer inspection, the distributions of the z-scored responses for both NPI conditions appear to be bimodal, as illustrated in Figure 6. This is further confirmed by Hall and York's critical bandwidth test (Hall and York, 2001), which resulted in critical bandwidths of 0.295 and 0.34467 for in-situ and ex-situ conditions, respectively. The associated *p*-values of .034 and .006 (respectively) indicate significant evidence against the null hypothesis of unimodality.

A by-item inspection did not flag any experimental items as abnormal, while a by-participant inspection showed a wider range of responses to NPI conditions compared to non-NPI conditions. In this section, we explore the possibility that the source of bimodality is different subpopulations of participants combined in a single sample. Our

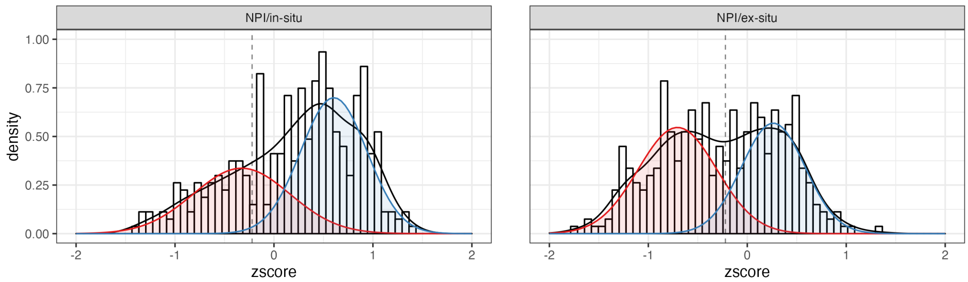


Fig. 6: Distributions of z-scored responses with overlaying mixture components and anti-mode cut points

working hypothesis is that one of the populations allows long-distance NPI licensing, while the other has a penalty associated with it.[22]

Since the experimental manipulation in Experiment 5 relies on long-distance NPI licensing (see Section 2.2 for details), if a participant's grammar does not permit it in the first place, this invalidates the experimental manipulation. Thus, the goal of the post hoc analysis is to identify participants whose grammar allows for long-distance NPIs. To do this, we applied cluster analysis methods to the z-scored responses for the NPI/in-situ control condition.

The clusterization solution identified approximately half of the participants in the original sample (which we refer to as "the first batch") as allowing long-distance NPIs. To maintain the same power as in Experiment 4, a second sample of the same size (i.e. "the second batch") was collected and the same exact processing and clustering steps were applied to it. Finally, the two groups of participants that allowed long-distance NPIs from both batches were combined and analyzed together.

4.4.1 The steps of the cluster analysis

The cluster analysis reported here consists of four steps: model selection, clustering, cluster validation, and grouping of participants. The analysis uses k-means clustering, an unsupervised learning method that partitions a dataset into k clusters.

[22]The results of the sanity check suggest that local NPI licensing remains unaffected and uniform.

Our working hypothesis states that there are several subpopulations in the sample, but it does not specify their number.[23] In order to select the model with the optimal number of clusters, we fit multiple finite Gaussian mixture models with different parameters to the experimental data using R packages `mclust` (Scrucca et al., 2016) and `factoextra` (Kassambara and Mundt, 2020). Since the clusterization data has one dimension (z-scores), we only considered equal-variance (E) and variable-variance (V) models. The comparison included all E and V models with the number of components G between 1 and 9. The goodness-of-fit was evaluated by computing the Bayesian Information Criterion (BIC) for each model, which takes into account both the likelihood of the data and the number of parameters in the model.[24] As a reminder, lower BIC values indicate a better fit. At the same time, negative BIC is often used for visualization purposes, since it places the winning model at the top.

During the clustering step, we applied the k-means clustering algorithm with the optimal number of clusters to the dataset. This algorithm minimizes the sum of squared distances between the data points and the closest cluster centroids. It is sensitive to the choice of initial centroids, which are selected randomly. Because of this, the clustering step was repeated 100000 times and the best solution was chosen based on the highest average silhouette width of a clustering solution. Furthermore, the package `mclust` provides model-based uncertainty estimates, such as the average posterior probabilities of cluster membership, which can be used to assess the stability of cluster solutions and the confidence in assigning individual observations to specific clusters.

[23] A reasonable possibility is that there are only two subpopulations of participants in the sample, those that allow long-distance NPI licensing and those that do not, but one can easily imagine a more gradient scenario requiring a partition with a larger number of k clusters.

[24] We chose to use BIC at this step instead of more familiar Akaike Information Criterion (AIC) because BIC incorporates a larger penalty for the number of parameters in the model compared to AIC, which can help prevent overfitting and improve the generalizability of the selected model. Another common alternative to AIC and BIC is the integrated complete-data likelihood (ICL) criterion (Biernacki et al., 2000). However, ICL is more appropriate for cluster spaces with greater separation between groups, while BIC provides a good conservative estimate of the number of components needed to approximate a density function in any cluster space.

Next, during the cluster validation step, we used two cluster validation methods. One is the silhouette plots, which were created using the `factoextra` package. Silhouette plots are a visual tool that shows the similarity (represented as silhouette width) between a data point and other points in the same cluster compared to other clusters. The silhouette width for a data point ranges between -1 and 1, where a high value indicates that the data point is well-matched to its own cluster, and a low value suggests that the data point is not well-matched. An average silhouette width greater than 0.6 is considered a good fit. As a second validation method, we calculated the Davies-Bouldin index (DBI; Davies and Bouldin 1979) using the `clusterSim` package (Walesiak and Dudek, 2020). DBI is defined as the average similarity across pairs of clusters, where similarity is the Euclidean distance between the centroids of two clusters divided by the scatter of the data points in the same cluster relative to its centroid. DBI is symmetric and nonnegative and ranges from 0 to ∞, with 0 indicating a perfect clustering solution, where each cluster is robust and far from others. A DBI value less than 0.8 is considered good, while a value > 1 suggests poor clustering.

Lastly, individual responses with their assigned cluster values were linked back to the participants. Based on their responses, the participants were assigned to different groups. During the experiment, each participant gave two responses per experimental condition. Therefore, the number of groups of participants can be calculated as the number of combinations with repetitions for the number of clusters k and the number of observations $o = 2$ using the following formula:

(42)

$$\frac{(k + o - 1)!}{o!(k - 1)!}$$

4.4.2 Clusterization results for the first batch

Figure 7 shows the results of each step of the cluster analysis applied to the first batch of participants in Experiment 5. At the model selection step, the BIC value suggests

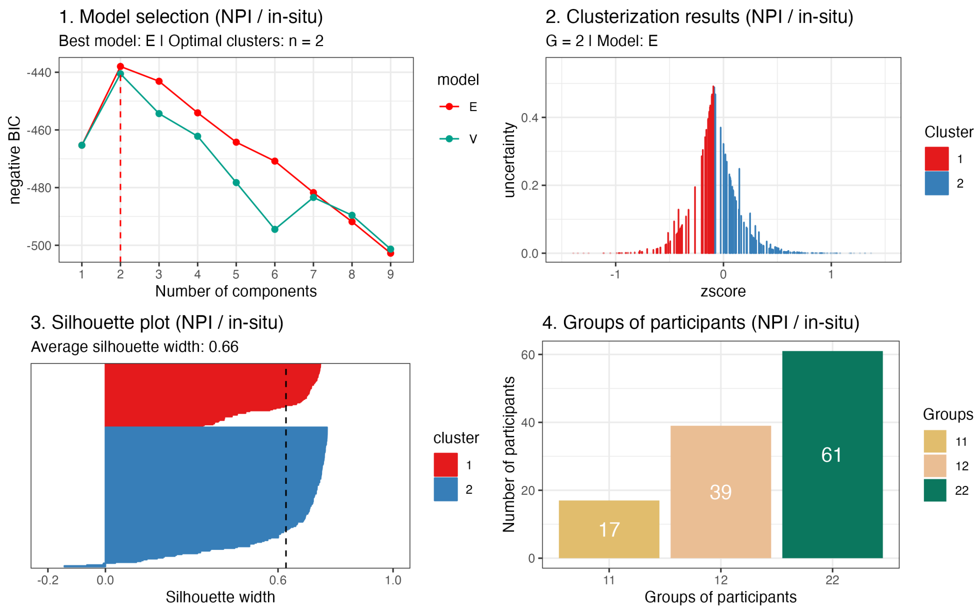


Fig. 7: The steps of the cluster analysis of the first batch (n=117)

that the best model for the data is an equal-variance mixture model with 2 clusters. The best clusterization solution for the model with these parameters is shown in Step 2. The amount of uncertainty in the central part indicates that the two clusters are close to each other. The first cluster corresponds to the responses that penalize and possibly disallow long-distance NPI licensing, while the second cluster contains responses consistent with long-distance NPI licensing being part of the participant's grammar. The silhouette plot in Step 3 suggests that the clustering solution is good enough (the average silhouette width > 0.6). The DBI value of 0.6 supports this conclusion. However, the silhouette plot also shows some negative width values in the second cluster, suggesting that the second cluster may be inflated with items misattributed from the first cluster. In the next step, linking the cluster values of the responses back to participants splits the participants into three groups: those who gave two responses from the first cluster (11, "one-one"), two responses from the second cluster (22, "two-two"), and those who gave mixed responses (12, "one-two"). Only group 22 is kept for further analysis because, for each participant, both of their responses show that their grammar allows long-distance NPI licensing, which is a necessary assumption for the experimental manipulation. Groups 11 and 12 are discarded because the participants' responses there indicate that they penalize long-distance NPI licensing in some way.[25]

4.4.3 Clusterization results for the second batch

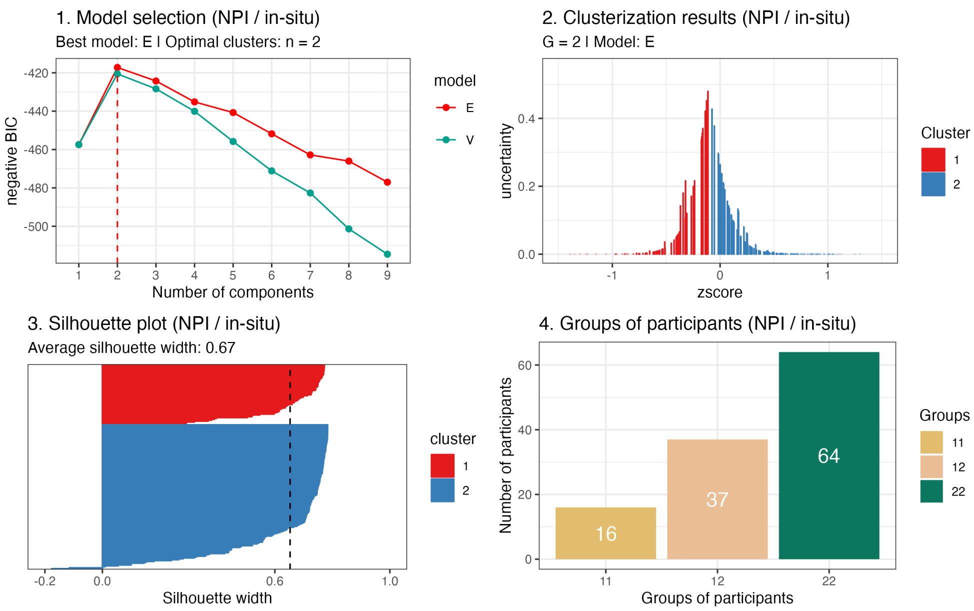


Fig. 8: The steps of the cluster analysis of the second batch (n=117)

The best model for the data according to the BIC value is a 2-cluster equal-variance Gaussian mixture model. The best clusterization solution is shown in Step 2. The first cluster contains responses that disallow long-distance NPI licensing, while the second cluster includes responses from individuals who have long-distance NPI licensing in their grammar. The silhouette plot in Step 3, with an average silhouette width greater than 0.6, indicates that the clustering solution is good. The DBI value

[25]The small spillover from cluster 2 into cluster 1 that is visible on the silhouette plot further suggests that the mixed response group 12 should be treated on a par with group 11 and not group 22.

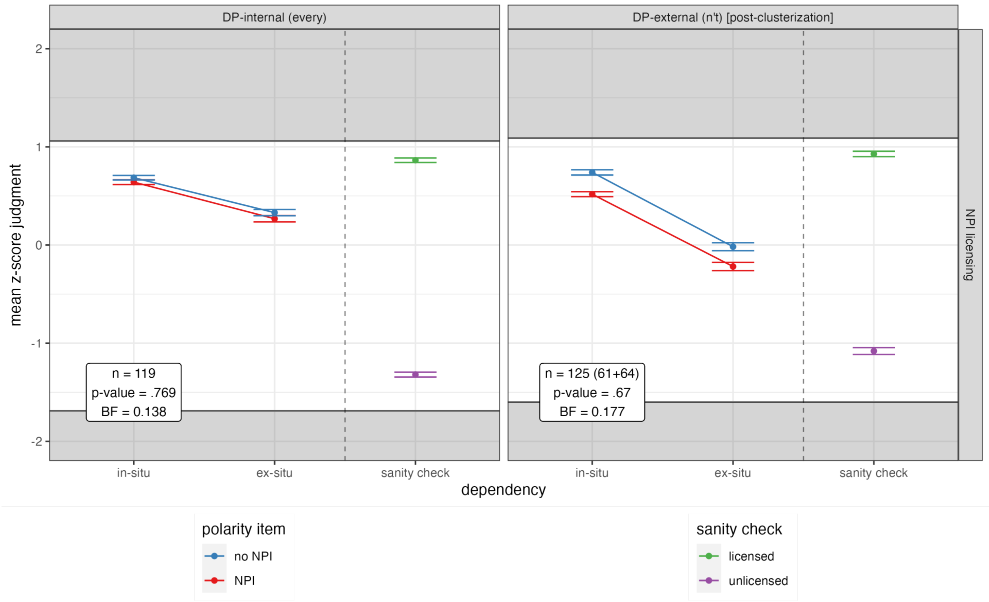
of 0.6 confirms this. However, the silhouette plot also reveals some negative width values in the second cluster, indicating that it may be saturated with misattributed items from the first cluster. Group 22 is kept for further analysis.

The results of the cluster analysis applied to the second batch of participants in Experiment 5 are shown in Figure 8. The best model for the data according to the BIC value is a 2-cluster equal-variance Gaussian mixture model. The best clusterization solution is shown in Step 2. The uncertainty estimates in the middle suggest that the two clusters are close to each other. Similarly to the first batch, the first cluster contains responses that disallow long-distance NPI licensing, while the second cluster includes responses from individuals who have long-distance NPI licensing in their grammar. The silhouette plot in Step 3, with an average silhouette width greater than 0.6, indicates that the clustering solution is good enough. This conclusion is further confirmed by the DBI value of 0.61. However, the silhouette plot also reveals some negative width values in the second cluster, indicating that it may be saturated with misattributed items from the first cluster. Linking the cluster attribution of individual responses back to the participants allows us to categorize them into the same three groups. Similar to the first batch, only group 22 is kept for further analysis.

4.4.4 A new analysis of Experiment 5 keeping only participants that allow long-distance NPI licensing

Using cluster analysis, we identified two groups of participants that allow long-distance NPI licensing in Experiment 5. Both groups were combined before performing the same statistical analysis again. Figure 9 contains the results of both NPI experiments, but shows the updated results of Experiment 5. Both sets of statistical tests indicate that there is no evidence of superadditive interaction in either experiment and, moreover, $BF_{10} < 0.33$ provides strong evidence for its absence in both experiments,

Fig. 9: Interaction plots for both NPI experiments (after clusterization)



suggesting that NPI licensing is not affected by RC extraposition in both DP-internal and DP-external configurations.

4.5 Discussion

The results of NPI licensing experiments are summarized in Table 4. We observe that RC extraposition does not interrupt NPI licensing in both DP-internal and DP-external configurations. Unlike in Condition C experiments, RC extraposition here shows uniform results, suggesting that a copy of the extraposed RC appears in the scope of both downward-entailing operators when NPI licensing is evaluated.

Table 4: Summary of experimental results testing the effect of RC extraposition on NPI licensing in English

#	C-command diagnostic	Syntactic configuration	Result
4	NPI licensing	DP-internal	licensed
5	NPI licensing	DP-external	licensed

5 General discussion

Table 5 consolidates the results of the Condition C and NPI licensing experiments. RC extraposition shows the obligatory reconstruction pattern in the DP-internal configuration and the optional reconstruction pattern in the DP-external configuration. The literature offers two explanations for the first result: either the RC moves to the right by itself and then obligatorily reconstructs to its base position, in line with Ross (1967), or the host of the RC is part of the target of RC extraposition and therefore appears in both the base and the extraposed position of the RC (Fox and Nissenbaum, 1999, 2000; Sportiche, 2016, 2019). However, the DP-external results exclude one of the interpretations: if the RC were to be obligatorily reconstructed to its base position, we would expect a Condition C violation in both configurations. Therefore, we can conclude that the target of RC extraposition includes both the RC and its host.

Table 5: Summary of experimental results testing the effect of RC extraposition on NPI licensing and Condition C in English

#	C-command diagnostic	Syntactic configuration	Result
1	Condition C	DP-internal	violation
2	Condition C	DP-external	no violation
4	NPI licensing	DP-internal	licensed
5	NPI licensing	DP-external	licensed

Table 6 repeats the predictions made by all possible combinations of a rightward movement theory of RC extraposition and a reconstruction pattern. All theories that accurately predict empirical data are color-coded. Notably, we have invalidated Ross's surface-friendly theory of RC extraposition as none of its combinations with reconstruction patterns yield correct predictions. Next, the obligatory LM version of the QR+LM theory has to adopt semantic reconstruction for NPI licensing. Therefore, contrary to Fox (2002), a Late-Merged RC needs to be able to undergo semantic reconstruction to the base position of its host. The remaining two theories of RC extraposition need to assume that an extraposed RC either can optionally reconstruct or is unable to reconstruct except via semantic reconstruction. Overall, the results of our study have significantly narrowed the space of possible theories of RC extraposition.

Table 6: Comparing predictions of RC extraposition theories for Condition C and NPI licensing across DP-internal and DP-external configurations with empirical results

Theory	Condition C		NPI licensing	
	DP-internal	DP-external	DP-internal	DP-external
Ross (1967)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	no violation	no violation	licensed	licensed
+ anti-reconstruction	no violation	no violation	unlicensed	unlicensed
+ semantic reconstruction	no violation	no violation	licensed	licensed
Fox and Nissenbaum (1999, 2000) + Obligatory LM (Stepanov, 2001b; Abe, 2018; Zyman, 2022)				
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed
Fox and Nissenbaum (1999, 2000) + Optional LM (Lebeaux, 1988)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	violation	no violation	licensed	licensed
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed
Sportiche (2016, 2019)				
+ obligatory reconstruction	violation	violation	licensed	licensed
+ optional reconstruction	violation	no violation	licensed	licensed
+ anti-reconstruction	violation	no violation	licensed	unlicensed
+ semantic reconstruction	violation	no violation	licensed	licensed

6 Conclusion

The general goal of this study was to determine the size of the constituent that undergoes RC extraposition in English. To do this, we conducted four experiments testing the effect of RC extraposition on Condition C violations and NPI licensing across two syntactic configurations. Our findings identified a discrepancy between the reconstruction patterns of RC extraposition in the DP-internal and DP-external configurations. RC extraposition shows obligatory reconstruction in the DP-internal configuration but optional reconstruction in the DP-external one. The only way to reconcile these results is by assuming that the constituent that undergoes RC extraposition is larger than the RC itself and includes the rest of the DP. This conclusion raises an interesting question about the place of RC extraposition among other rightward movement phenomena. If RC extraposition is an instance of the rightward DP movement, it should be analyzed on a par with its other instances such as Stylistic and Locative Inversion and (possibly) Heavy NP Shift. We believe that selecting the right comparison class for RC extraposition may be the first step towards understanding some of its otherwise puzzling syntactic and semantic properties (Perlmutter and Ross, 1970; Link, 1984; Baltin, 1981, 1987; Culicover and Rochemont, 1990; Huck and Na, 1990).

References

Abe, J. 2018. Late Merge and Phases for Anti-C-Command Requirements. Syntax 21(2): 91–111 .

Abney, S.P. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Al-Aqarbeh, R. and J. Sprouse 2022. Island effects and amelioration by resumption in Jordanian Arabic: an auditory acceptability judgment study. Manuscript.

Baltin, M. 1981. Strict bounding, In The Logical Problem of Language Acquisition, eds. Baker, C. and J. McCarthy, 257–295. The MIT Press.

Baltin, M.R. 1987. Do antecedent-contained deletions exist? Linguistic Inquiry 18(4): 579–595 .

Barker, C. 2012. Quantificational binding does not require c-command. Linguistic inquiry 43(4): 614–633 .

Barss, A. and H. Lasnik. 1986. A Note on Anaphora and Double Objects. Linguistic inquiry 17(2): 347–354 .

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models using lme4. Journal of Statistical Software 67(1): 1–48. https://doi.org/10.18637/jss.v067.i01 .

Biernacki, C., G. Celeux, and G. Govaert. 2000. Assessing a mixture model for clustering with the integrated completed likelihood. IEEE transactions on pattern analysis and machine intelligence 22(7): 719–725 .

Bruening, B. 2018. Double object constructions and prepositional dative constructions are distinct: A reply to Ormazabal and Romero 2012. Linguistic Inquiry 49(1): 123–150 .

Chomsky, N. 1981. Lectures on Government and Binding. Foris.

Chomsky, N. 1993. A minimalist program for linguistic theory, In The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger, eds. Hale, K. and S.J. Keyser, 1–52. The MIT Press.

Corver, N. and J. Nunes. 2007. The copy theory of movement. Amsterdam/Philadelphia: John Benjamins.

Culicover, P. and M. Rochemont. 1990. Extraposition and the complement principle. Linguistic Inquiry 21(1): 23–47 .

Davies, D.L. and D.W. Bouldin. 1979. A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence PAMI-1(2): 224–227 .

Dayal, V. 2009. Variation in English free choice items. Universals and variation: Proceedings of GLOW in Asia VII: 237–256 .

Fox, D. 1999. Reconstruction, binding theory, and the interpretation of chains. Linguistic inquiry 30(2): 157–196 .

Fox, D. 2002. Antecedent-contained deletion and the copy theory of movement. Linguistic inquiry 33(1): 63–96 .

Fox, D. 2017a. Embedded Late Merge and the theory of displacement. Manuscript.

Fox, D. 2017b. Non-local extraposition: a modification of Williams's Generalization, In A Schrift to Fest Kyle Johnson, eds. LaCara, N., K. Moulton, and A.M. Tessier, 131–139. Linguistics Open Access Publications 1.

Fox, D. and K. Johnson 2016. QR is restrictor sharing. In Proceedings of the 33rd West Coast Conference on Formal Linguistics. Somerville, Mass.: Cascadilla Proceedings Project, pp. 1–16.

Fox, D. and J. Nissenbaum 1999. Extraposition and Scope: A case for overt QR. In S. Bird (Ed.), Proceedings of the 18th West Coast Conference in Formal Llinguistics, pp. 132–144.

Fox, D. and J. Nissenbaum. 2000. Extraposition and covert movement. Talk presented at NYU .

Freidin, R. 1986. Fundamental Issues in the Theory of Binding, In Studies in the Acquisition of Anaphora, ed. Lust, B., Number 2 in Studies in Theoretical Psycholinguistics. Springer.

Fukuda, S., N. Tanaka, H. Ono, and J. Sprouse. 2022. An experimental reassessment of complex NP islands with NP-scrambling in Japanese. Glossa: a journal of general linguistics 7(1) .

Gajewski, J. 2016. Another look at NPIs in definite descriptions: An experimental approach, Negation and polarity: experimental perspectives, 307–327. Springer.

Giannakidou, A. 2002. Licensing and sensitivity in polarity items: from downward entailment to nonveridicality. CLS 38: 29–53 .

Gordon, P.C. and R. Hendrick. 1997. Intuitive knowledge of linguistic co-reference. Cognition 62(3): 325–370 .

Guerzoni, E. and Y. Sharvit. 2007. A question of strength: on NPIs in interrogative clauses. Linguistics and Philosophy 30(3): 361–391 .

Hall, P. and M. York. 2001. On the calibration of silverman's test for multimodality. Statistica Sinica: 515–536 .

Harley, H. and S. Miyagawa. 2018. Syntax of ditransitives, Oxford Research Encyclopedia of Linguistics. Oxford University Press.

Heycock, C. 1995. Asymmetries in reconstruction. Linguistic inquiry: 547–570 .

Hoeksema, J. 1986. Monotonicity phenomena in natural language, Volume 5. Semantics–Critical Concepts in Linguistics.

Hoeksema, J. 2017. Neg-raising and long-distance licensing of negative polarity items. Negation and contact: With special focus on Singapore English. Amsterdam & Philadelphia: John Benjamins: 33–61 .

Homer, V. 2010. Polarity and modality. Doctoral dissertation, University of California, Los Angeles, Los Angeles, CA.

Horn, L.R. 1989. A Natural History of Negation. University of Chicago Press.

Huck, G.J. and Y. Na. 1990. Extraposition and focus. Language: 51–77 .

Hulsey, S. and U. Sauerland. 2006. Sorting out relative clauses. Natural language semantics 14(2): 111–137 .

Iglewicz, B. and D.C. Hoaglin. 1993. How to detect and handle outliers, Volume 16. Asq Press.

Jeffreys, H. 1939. The theory of probability. Oxford University Press.

Johnson, K. 2020. Rethinking linearization, In Syntactic architecture and its consequences II: Between syntax and morphology, eds. Bárány, A., T. Biberauer, J. Douglas, and S. Vikner, Number 10 in Open Generative Syntax, 113–136. Language Science Press.

Kassambara, A. and F. Mundt 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7.

Kazanina, N., E.F. Lau, M. Lieberman, M. Yoshida, and C. Phillips. 2007. The effect of syntactic constraints on the processing of backwards anaphora. Journal of Memory and Language 56(3): 384–409 .

Keller, F. and A. Asudeh 2001. Constraints on linguistic coreference: Structural vs. pragmatic factors. In Proceedings of the Annual Meeting of the Cognitive Science Society, Volume 23.

Klima, E.S. 1964a. Negation in English, In The structure of language, eds. Fodor, J.A. and J.J. Katz. Prentice-Hall.

Klima, E.S. 1964b. Relatedness between grammatical systems. Language 40(1): 1–20 .

Koval, P. and J. Sprouse 2023. Island sensitivity of relative clause extraposition: Evidence for rightward movement. Manuscript.

Kuznetsova, A., P.B. Brockhoff, and R.H.B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software 82(13): 1–26 .

Ladusaw, W.A. 1979. Polarity Sensitivity as Inherent Scope Relations. The University of Texas at Austin.

Larson, R.K. 1988. On the double object construction. Linguistic inquiry 19(3): 335–391 .

Lasnik, H. 1989. Essays on anaphora, Volume 16 of Studies in Natural Language and Linguistic Theory. Kluwer.

Lebeaux, D. 1988. Language Acquisition and the form of grammar. Doctoral dissertation, University of Massachusetts Amherst, Amherst, MA.

Lebeaux, D. 2000. Language acquisition and the form of the grammar. John Benjamins Publishing.

Lebeaux, D. 2009. Where does binding theory apply? MIT Press.

Linebarger, M.C. 1980. The grammar of negative polarity. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Link, G. 1984. Hydras: On the logic of relative constructions with multiple heads, In Varieties of Formal Semantics, eds. Landman, F. and F. Veltman, 151–180. Reidel.

Mahajan, A.K. 1990. LF conditions on negative polarity licensing. Lingua 80(4): 333–348 .

Marty, P., E. Chemla, and J. Sprouse. 2020. The effect of three basic task features on the sensitivity of acceptability judgment tasks. Glossa: a journal of general linguistics 5(1): 72 .

Morey, R.D. and J.N. Rouder 2018. BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2.

Moulton, K. 2013. Not moving clauses: Connectivity in clausal arguments. Syntax 16(3): 250–291 .

Neyman, J. and E.S. Pearson. 1928. On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. Biometrika 20(1/2): 175–240 .

Overfelt, J. 2015. Extraposition of NPIs from NP. Lingua 164: 25–44 .

Partee, B. 1975. Montague grammar and transformational grammar. Linguistic inquiry: 203–300 .

Perlmutter, D.M. and J.R. Ross. 1970. Relative clauses with split antecedents. Linguistic Inquiry: 350 .

Progovac, L. 1988. A Binding Approach to Polarity Sensitivity. Doctoral dissertation, University of Southern California, Los Angeles, CA.

Progovac, L. 1993. Negative polarity: Entailment and binding. Linguistics and Philosophy 16(2): 149–180 .

R Core Team 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Reinhart, T. 1976. The syntactic domain of anaphora. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Reinhart, T. 1981. Definite NP anaphora and c-command domains. Linguistic inquiry 12(4): 605–635 .

Romero, M. 1998. Focus and reconstruction effects in wh-phrases. Doctoral dissertation, University of Massachusetts Amherst, Amherst, MA.

Ross, J.R. 1967. Constraints on Variables in Syntax. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Scrucca, L., M. Fop, T.B. Murphy, and A.E. Raftery. 2016. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. The R Journal 8(1): 289–317 .

Sportiche, D. 2016. Neglect. Manuscript.

Sportiche, D. 2017. Relative clauses: Promotion only, in steps. Manuscript available at http://ling.auf.net/lingbuzz/003444.

Sportiche, D. 2019. Somber prospects for late merger. Linguistic Inquiry 50(2): 416–424 .

Sprouse, J. and D. Almeida. 2017. Design sensitivity and statistical power in acceptability judgment experiments. Glossa 2(1): 1 .

Sprouse, J., C.T. Schütze, and D. Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. Lingua 134: 219–248 .

Stepanov, A. 2001a. Cyclic domains in syntactic theory. Doctoral dissertation, University of Connecticut, Storrs, CT.

Stepanov, A. 2001b. Late adjunction and minimalist phrase structure. Syntax 4(2): 94–125 .

Stockwell, R., A. Meltzer-Asscher, and D. Sportiche. 2021. There is reconstruction for Condition C in English questions. North East Linguistic Society (NELS 51) .

Stockwell, R.P., P. Schacter, and B.H. Partee. 1973. The major syntactic structures of English. Holt, Rinehart and Winston.

Takahashi, S. and S. Hulsey. 2009. Wholesale late merger: Beyond the A/ā distinction. Linguistic Inquiry 40(3): 387–426 .

Temme, A. and E. Verhoeven. 2017. Backward binding as a psych effect: A binding illusion? Zeitschrift für Sprachwissenschaft 36(2): 279–308 .

Tukey, J.W. 1977. Exploratory Data Analysis. Mass.: Reading.

Uribe-Etxebarria, M. 1995. Negative polarity item licensing, indefinites and complex predicates. In Semantics and Linguistic Theory, Volume 5, pp. 346–361.

von Fintel, K. 1999. NPI licensing, Strawson entailment, and context dependency. Journal of semantics 16(2): 97–148 .

Walesiak, M. and A. Dudek 2020. The Choice of Variable Normalization Method in Cluster Analysis. In K. S. Soliman (Ed.), Education Excellence and Innovation Management: A 2025 Vision to Sustain Economic Development During Global Challenges, pp. 325–340. International Business Information Management Association (IBIMA).

Wilder, C. 1996. Rightward movement as leftward deletion, On extraction and extraposition in German, Number 11 in Linguistik Aktuell / Linguistics Today, 273–309. John Benjamins Amsterdam.

Zyman, E. 2022. Phase-Constrained Obligatory Late Adjunction. Syntax 25(1): 84–121 .