

Accepted for publication in:

Competing Motivations in Grammar, Acquisition, and Usage, ed. by Edith Moravcsik, Andrej Malchukov, and Brian MacWhinney. Oxford University Press.

Please do not cite without permission.

Comments welcome and appreciated!

Why move? How weight and discourse factors combine to predict
relative clause extraposition in English

Elaine J. Francis
Purdue University
ejfranci@purdue.edu

Laura A. Michaelis
University of Colorado at Boulder
Laura.Michaelis@Colorado.EDU

Why move? How weight and discourse factors combine to predict
relative clause extraposition in English

Elaine J. Francis

Purdue University

Laura A. Michaelis

University of Colorado at Boulder

Abstract

In relative clause extraposition (RCE) in English, a subject-modifying relative clause appears after the verb phrase, as in *Some research was conducted that supports the existing theory*. Previous studies have revealed that both grammatical weight and discourse factors induce speakers to use RCE. However, the current study is the first to examine the interaction of the two. A quantitative analysis of RCE and comparable non-RCE tokens in the International Corpus of English Great Britain (ICE-GB) showed a strong effect of weight: there is a strong preference for RCE when the relative clause is at least five times longer than the verb phrase, and a strong preference for canonical (non-RCE) order when the relative clause is the same length or shorter than the verb phrase. However, for those tokens with length ratios falling between these limits, choice of structure depends primarily on discourse factors.

X.1 Introduction

In this chapter, we examine a well-known constituent-order option of English, relative clause

extraposition (RCE), and ask whether it subserves sentence processing, information packaging or both. In RCE from subject NP position (referred to by Ross 1967 as Extraposition from NP), a subject-modifying relative clause occurs following the VP, as in the example from the International Corpus of English Great Britain (henceforth ICE-GB) in (1a)¹, rather than adjacent to its head noun, as in the corresponding non-RCE sentence (1b).

- (1) a. Further research has been conducted on this that indicates this criticism may not be just. (ICE-GB)
- b. Further research that indicates this criticism may not be just has been conducted on this.

The alternative sentences in (1a-b) seem to express exactly the same meaning. Structurally, however, (1a) is more complex because there is a discontinuous dependency between the subject NP and the relative clause. Most syntactic accounts of RCE have avoided positing an actual discontinuous constituent, instead licensing RCE through rightward movement (Ross 1967, Baltin 1981), leftward movement (Kayne 1994), adjunction and co-indexing (Culicover and Rochemont 1990) or percolation of a list-valued EXTRA feature (Kay and Sag 2012). Nevertheless, this kind of dependency relation strains rule-to-rule semantic composition, while also requiring greater processing effort in word-by-word reading (Levy et al 2012). Further, unlike many other constructions featuring non-canonical word order, like wh-movement and topicalization, RCE has no obvious functional motivation.

Given the added complexity of RCE, and its apparent lack of semantico-pragmatic effect, one can reasonably ask why speakers use the construction. The literature has provided two main

answers. The first answer, provided by the majority of the studies, is that RCE is used to place presentational focus on the denotatum of the subject NP, and thus has a function similar to that of the presentational *there*-construction (PTC):

- (2) There exists further research on this that indicates this criticism may not be just.

The general consensus is that RCE is used to highlight new, contrastive or important information conveyed by the subject NP while backgrounding the information contained in the main-clause predicate (Huck and Na 1990; Kuno and Takami 2004; Rochemont and Culicover 1990; Takami 1999). Both the RCE sentence (1a) and the PTC sentence (2) assert that research findings of a particular type exist, rather than that research of this type was conducted.

Such a discourse-based approach has been used to explain some of the formal properties that are typically, but not categorically, associated with RCE. For example, the subject of an RCE sentence is typically an indefinite NP, reflecting the focal status of the NP's referent. However, focal status does not require indefiniteness; an RCE token may, for example, contain a definite subject NP interpreted as a contrastive focus (Huck and Na 1990). In (3), the speaker has already introduced two guys, and is contrasting the guy from Treno's with the guy from a different restaurant:

- (3) The guy just came in that I met at TRENO'S yesterday. (Huck and Na 1990: 54)

Similarly, RCE tokens typically contain a presentative matrix predicate. Following Levy et al. (2012: 17), we define a presentative predicate as one containing an intransitive head verb of

existence or appearance; such predicates are typically used to introduce the referent of the subject NP into a scene (e.g. *came in* in (3) above). However, Rochemont and Culicover (1990: 65-68) point out that it is possible to use a non-presentative verb in an RCE predication, provided that the predicate is “directly c-construable,”—already in the discourse (Rochemont 1986: 174). For example, the verb *scream* can be used when the property of screaming is already evoked in the discourse, as in (4):

(4) Suddenly there was the sound of lions growling. Several women screamed. Then a man screamed who was standing at the very edge of the crowd.

(Rochemont and Culicover 1990: 65)

In short, these discourse-based theories predict that definiteness and predicate type may vary among different RCE sentences while presentational or contrastive focus on the subject NP should occur consistently.

While these predictions are plausible, empirical support for them has mostly been provided by constructed examples (as in 3-4 above) rather than analysis of actual language use. One exception is Francis (2010), which did not investigate discourse status, but did look at corpus frequencies for different predicate types and found that although RCE occurred with presentative verbs more frequently than with transitive or non-presentative intransitive verbs, RCE was most commonly used with passive verbs (as in (1a) above). Arguing that passive verbs are semantically similar to presentative verbs (as intransitive predicates which select a Theme argument) and just as felicitous in presentational contexts, Francis (2010: 63) concluded that the

prevalence of passive predicates is compatible with theories that posit a presentational function for RCE.

The second functional rationale for the existence of RCE involves grammatical weight. Arnold and Wasow (2000), Wasow (2002) and Hawkins (2004), among others, have shown that shifting heavy (i.e. long and/or syntactically complex) constituents to the end of a clause can facilitate language production and comprehension. In production, heavy constituents may be difficult to formulate. Therefore, postponing them offers speakers more time to finish formulating the sentence while they produce the shorter, easier phrases (Arnold and Wasow 2000: 32). In comprehension, shifting heavy constituents to the end allows listeners to reduce the storage and integration costs associated with non-local dependencies (Gibson 1998; Hawkins 2004). For example, although RCE always increases the distance between the head noun and its relative-clause modifier, a long relative clause in a non-RCE sentence increases the distance between the subject NP and its predicate. Hawkins' (2004) theory of domain minimization, which quantifies the notion of integration distance in terms of cumulative effects across different domains, thus predicts a greater overall processing cost for the non-RCE sentence (5b) as compared with the RCE sentence (5a). This is because the relative clause which separates the subject from the predicate in (5b) is much longer than the VP which separates the head noun from the relative clause in (5a).

- (5) a. Certain conditions existed which cannot be applied to all other countries at all times.
(ICE-GB)
- b. Certain conditions which cannot be applied to all other countries at all times existed.

Consistent with these predictions, Francis (2010) found significant effects of grammatical weight in both production and comprehension. An analysis of RCE and non-RCE tokens in the ICE-GB corpus showed that RCE was strongly preferred over canonical (non-RCE) order when the relative clause was at least five times longer (in words) than the main-clause VP, whereas RCE occurred only rarely when the VP was the same length or longer than the relative clause. In addition, a full-sentence reading-time task showed that RCE sentences were read significantly faster than non-RCE sentences when the relative clause was long, but that there was no difference in reading time with short relative clauses.²

As these previous studies have shown, both discourse factors and grammatical weight appear to play a role in speakers' use of RCE. However, no previous studies have examined both factors simultaneously. Thus, it is not known to what extent discourse factors are independent of weight, nor is it known which factors have the strongest influence over speakers' choice of structure. To address these questions, we conducted a quantitative analysis of naturally occurring tokens of RCE and non-RCE from the International Corpus of English-Great Britain (ICE-GB). Our findings confirmed that RCE is in fact typically associated with presentational contexts, and that grammatical weight and certain discourse-related factors independently influence speakers' choice of structure.

Although this study is the first of its kind to examine RCE in English, it aligns with several other recent studies of constituent-order alternations in English, including the ditransitive alternation (Bresnan and Ford 2010), particle shift (Gries 2003; Lohse et al. 2004), genitive placement (Rosenbach 2005), heavy NP shift (Arnold et al. 2000), and object fronting in Dutch (Lamers and de Hoop, this volume)—as well as the closely-related RCE construction in German (Strunk, this volume). These studies have all shown that constituent-order alternations tend to be

conditioned by multiple interacting factors, including grammatical weight, information structure, animacy, lexical bias and structural priming.

The remainder of this chapter is organized as follows. Section X.2 describes the methods and quantitative results of a corpus analysis of RCE and non-RCE tokens in the ICE-GB corpus. Section X.3 presents a qualitative analysis of exceptional cases from the corpus. Finally, Section X.4 briefly outlines some implications of the current study for linguistic theory.

X.2 Corpus study

ICE-GB (Nelson et al. 2002) includes about one million words of British English, culled from a variety of spoken and written genres; it is parsed and tagged to facilitate identification of syntactic structures. The current analysis is based on a subset of sentences from the ICE-GB, which were originally collected for Francis (2010). Using Francis' (2010) original coding for phrase length and predicate type, the current analysis provides additional coding for several discourse-related categories. The remainder of this section describes the coding scheme, hypotheses and quantitative results.

X.2.1 Coding scheme

This analysis includes 345 sentences with a lexical subject NP modified by a finite relative clause—the total number of such sentences found in the ICE-GB corpus. These were collected as a subset of the 391 sentences from Francis (2010), excluding the 46 tokens with pronominally headed relative clauses. Pronominally headed relative clauses were excluded from the present analysis because they are known to have somewhat different information-structure properties from lexically headed RCs, and we wanted to eliminate variation due to the activation status of

the nominal-head denotatum (Gundel et al. 1993; Michaelis and H. Francis 2007). These sentences were originally extracted from the corpus using tree-fragment searches to identify the non-RCE tokens and function searches (specifically, “floating post-nominal modifier”) in combination with manual checking to identify the RCE tokens. Grammatical weight was coded according to the original measurements in Francis (2010): VP length (in words), RC length (in words), and VP-to-RC length ratio (VP length divided by RC length).³ VP-to-RC length ratio was used as a measure of relative length, in accordance with previous corpus and experimental studies that have found relative length to be a more significant predictor of word-order choice than absolute length (Hawkins 1994; Stallings and MacDonald 2011; Wasow and Arnold 2003). Length difference (VP length minus RC length) was calculated as an additional measure of relative length, to determine which measure made the more accurate predictions. We used the predicate type coding from Francis (2010). Specifically, we distinguished between passive or presentative predicates and other predicate types. All morphologically passive verbs (consisting of *be* + past participle) were counted in the passive category. Following Rochemont and Culicover (1990: 66), who claim that only presentative predicates are felicitous with PTC (whereas RCE allows a wider range of predicates), we operationally defined presentative predicates as those active intransitive verbs that remain felicitous when the sentence is converted from an RCE or non-RCE sentence, as in (6a-b), to a PTC sentence, as in (6c):

- (6) a. Certain conditions existed which cannot be applied to all other countries at all times.
(ICE-GB)
- b. Certain conditions which cannot be applied to all other countries at all times existed.

c. There existed certain conditions which cannot be applied to all other countries at all times

As Ward and Birner (1996: 469-471) show based on 428 corpus examples, PTC sentences appear to require a post-verbal NP denoting a discourse-new entity. Because we wanted our presentative category to be independent from definiteness and discourse status of the subject NP referent, we also counted a predicate as presentative if it became felicitous with PTC when the NP was changed from definite to indefinite.

In addition, we coded each sentence for definiteness of the main-clause subject NP, discourse status of the main-clause subject NP, and discourse status of the main-clause VP. Subject NPs with a definite article (*the*), demonstrative determiner (*this, that, these, those*), strong quantifier (*all, both, each, every, most*), or possessive determiner (*our, your, his, etc.*) were classified as definite, while subject NPs with an indefinite article (*a, an*), weak quantifier (*some, many, few, no, several, one, cardinal numbers*), or no determiner (e.g., *people*) were classified as indefinite, based on Carlson's 1977 claim that bare nouns are intrinsically non-quantificational. Classifications of discourse status were based on the preceding 20 lines of text. Following Michaelis and H. Francis (2007) and Gregory and Michaelis (2001), we used three categories to label information status: given (prior mention), superset mention, and new (no prior mention). Subject NPs with a prior mention of the same referent within the preceding 20 lines were classified as given. Subject NPs with a prior mention of the category including the referent but no prior mention of the referent itself were classified as superset mention (Michaelis and H. Francis 2007: 28). For example, for the sentence *The point that Paula made was well justified*, the subject NP would be classified as superset mention if there were a prior mention of a point

made by another speaker, or if the general idea of making a point had been brought up. Subject NPs with no prior mention of the referent or of the category including the referent were classified as new. Predicate VPs were coded in a similar manner. Predicate VPs with a prior mention of the exact same event/situation within the preceding 20 lines of text were classified as given. Predicate VPs with a prior mention of the type of event/situation (but not the exact event/situation) were classified as superset mention. For example, for the sentence *The organ which you hear is over 100 years old*, the predicate VP would be classified as superset mention if there were a previous discussion of something else (not this particular organ) being over 100 years old. Predicate VPs with no prior mention of the exact event/situation or of the type of event/situation denoted by the predicate were classified as new.

Two independent raters classified every item for the discourse status of the subject NP and the predicate VP. A third rater then independently rated all of the items for which the first two raters disagreed, and for those cases, the category selected by two out of the three raters was used for the analysis. Overall, the first two raters agreed for 69% of the subjects and 89% of the predicates.

X.2.2 Discourse-related properties of RCE and non-RCE tokens

Although the main goal of this study is to determine whether and how grammatical weight and several discourse-related factors influence the speaker's choice of structure, we first briefly compare and contrast the respective discourse profiles of RCE and non-RCE clauses in order to verify whether our corpus data are in line with previous claims from the discourse literature. Owing to limitations of space, we include only descriptive statistics here.

As shown in Figure 1 and Table 1, RCE tokens had predominantly indefinite subject NPs,

with passive or presentative predicates. Non-RCE tokens differed from RCE tokens in having predominantly definite subject NPs and fewer passive and presentative predicates. Further analysis revealed that non-RCE tokens occurred predominantly with transitive and copular predicates.

Figure 1: Definiteness, predicate type, and discourse status for RCE tokens (n = 53) and non-RCE tokens (n = 292)

[Figure 1 here]

Table 1: Proportions and counts for RCE and non-RCE tokens with respect to definiteness and predicate type

	Definiteness		Predicate Type		
	% of RCE Tokens	% of Non-RCE Tokens	% of RCE Tokens	% of Non-RCE Tokens	
Definite	19% (n = 10)	61% (n = 178)	Passive/ Presentative	70% (n = 37)	21% (n = 61)
Indefinite	81% (n = 43)	39% (n = 114)	Other Predicate	30% (n = 16)	79% (n = 231)
Total	100% (n = 53)	100% (n = 292)	Total	100% (n = 53)	100% (n = 292)

As shown in Figure 1 and Table 2, RCE and non-RCE tokens differed relatively little with regard to discourse status: tokens of both kinds typically contained discourse-new subjects and discourse-new predicates.⁴ However, there were some differences. Non-RCE tokens more frequently contained discourse-given subject NPs than did non-RCE tokens (19.9% vs. 1.9%).

This finding is consistent with the traditional analysis of RCE as a presentational construction and with the analysis of non-RCE subjects as topics. However, it is notable that both RCE tokens and non-RCE tokens occurred predominantly with discourse-new subject NPs. This tendency is not predicted by previous accounts of RCE, but is in line with a corpus study by Michaelis and H. Francis (2007), which found that subject NPs headed by a lexical (common) noun were typically discourse new, despite expressing a sentence topic and containing a definite determiner. The authors proposed that sentences with a lexically headed subject NP typically serve a dual role of both introducing a new topic and commenting on it.

Table 2: Proportions and counts for discourse status of subject and predicate for RCE tokens (n = 53) and non-RCE tokens (n = 291)⁵

	Subject Accessibility		Predicate Accessibility	
	% of RCE Tokens	% of non-RCE Tokens	% of RCE Tokens	% of non-RCE Tokens
Given	1.9% (n = 1)	19.9% (n = 58)	1.9% (n = 1)	0.3% (n = 1)
Superset Mention	49.1% (n = 26)	29.6% (n = 86)	18.9% (n = 10)	3.1% (n = 9)
No Prior Mention	49.1% (n = 26)	50.5% (n = 147)	79.2% (n = 42)	96.6% (n = 281)
Total	100% (n = 53)	100% (n = 291)	100% (n = 53)	100% (n = 291)

RCE and non-RCE tokens did not differ in their dispreference for discourse-given predicates. However, breaking down the discourse-new category into “superset mention” and “no prior mention” reveals an interesting difference. Predicates of RCE tokens occur more often than predicates of non-RCE tokens in the superset mention category (18.9% vs. 3.1%), as shown

in Table 2. Thus, although predicates of RCE sentences were almost never discourse given, they did occur more often with a predicate that was accessible from the context.

In summary, the descriptive findings reported in this section appear to support a presentational analysis of RCE, as put forth in previous studies of extraposition (Huck and Na 1990; Kuno and Takami 2004; Rochemont and Culicover 1990). RCE tokens typically had indefinite, discourse-new subjects and passive or presentative predicates. In contrast, non-RCE tokens typically had definite, discourse-new subjects and transitive or copular predicates. The major unexpected finding was that the predicates of RCE tokens were almost never discourse-given. However, predicates of RCE tokens did occur in the superset mention category more often than predicates of non-RCE tokens. In fact, the lack of discourse-given predicates in our RCE data may be less surprising than it appears, if we view RCE predicates through the lens of informativeness rather than discourse status. Perhaps what distinguishes RCE predicates from those in canonical predications is the latter have relatively bleached semantics (e.g., denoting superordinate-level events or states), rather than ‘being in the presupposition’ (see Takami 1999: 27ff). We leave information-based labeling of predicates to a future effort.

X.2.3 Discourse and weight-based factors as independent predictors of RCE usage

In this section, we report findings related to the primary aim for this study—to determine the relative influence of grammatical weight and discourse factors on speakers’ choice of RCE as against a non-RCE structure. Although we will consider some of the same factors as in the previous section, we will begin to characterize these factors as *predictors* of structural choices. That is, definiteness and predicate type are among the independent variables that we will use to predict the value of the dependent variable, extraposition status (RCE or non-RCE). This

analysis is based on a binary logistic regression model. By entering all of the factors into a logistic regression model, we were able to determine whether each factor contributed independently to the predictive power of the model. In addition, the logistic regression model allowed us to test for interactions and to determine the relative strength of each factor.

Based on the descriptive results reported in the previous section, in which only two tokens had a discourse-given predicate, we framed our hypothesis regarding predicate accessibility with reference to the ‘superset mention’ category. To the discourse-based factors considered in the previous section we added the factor *length ratio* (a measure of grammatical weight). Our hypotheses were as follows:

- (1) Probability of RCE should be highest for tokens with the lowest VP-to-RC length ratio, and should decrease as this ratio increases (Francis 2010).
- (2) Probability of RCE should be higher for tokens with a passive or presentative predicate than for tokens with other predicate types (Francis 2010).
- (3) Probability of RCE should be higher for tokens with an indefinite subject NP than for tokens with a definite subject NP (Huck and Na 1990).
- (4) Probability of RCE should be higher for tokens with a superset-mention predicate than for tokens with a discourse-new (no prior mention) predicate (Rochemont and Culicover 1990).
- (5) Probability of RCE should be higher for tokens with a discourse-new subject (superset mention or no prior mention) than for tokens with a discourse-given subject (Rochemont and Culicover 1990).

We had no specific predictions regarding the relative strengths of the different factors, and no specific predictions regarding statistical interactions

Before reporting the results of the logistic regression analysis, we will first provide some basic descriptive statistics showing the trends in the data (Figures 2-6). In the discussion that follows, extraposition status will be expressed in terms of the percentage of all tokens from a particular category which had RCE word order.

Overall, canonical (non-RCE) word order was used much more frequently than RCE, with RCE used in only 15% of all tokens (53 of 345). However, frequency of RCE increased to 54% (31 of 57) for items that both had an indefinite subject NP and a passive or presentative predicate (Figure 2), while frequency remained at 15% or less for items belonging to the other three definiteness/predicate type combinations.

Figure 2: Percentage of RCE for four categories of definiteness and predicate type

[Figure 2 here]

Incidence of RCE also differed according to discourse status. As shown in Figure 3, RCE occurred very rarely in items with a given subject NP (1 of 59 tokens), but occurred more frequently in items with a superset-mention predicate (10 of 19 tokens). Consistent with the previous literature, the former trend suggests a dispreference for RCE outside of presentational contexts, while the latter trend suggests a slight preference for RCE when the main-clause predicate is relatively accessible. As discussed in X.2.2 above, there were only two tokens with discourse-given predicates in the entire sample. Thus, the apparently high rate of RCE shown in

Figure 3 (50%) represents one out of two tokens, and therefore does not indicate any identifiable trend.

Figure 3: Percentage of RCE by discourse status of subject and predicate

[Figure 3 here]

As in Francis (2010), dramatic differences in incidence of RCE were found for different ratios of VP to RC length. As shown in Figure 4, incidence of RCE was highest for length ratios of 0.2 (or 1:5) and lower (tokens for which the RC was at least five times longer than the VP), at 91%, and decreased as this ratio increased. For length ratios of 1.0 (or 5:5) and higher (tokens for which the RC was the same length or shorter than the VP), RCE occurred in only 2% of tokens.

Figure 4: Percentage of RCE for increasing ratios of VP length to RC length

[Figure 4 here]

Crucially, the effect of length ratio differed for items of different definiteness statuses and predicate types, as shown in Figures 5-6. For tokens with a definite subject NP (Figure 5), RCE was preferred over non-RCE order only when the VP-to-RC length ratio was less than 0.2, or 1:5 (i.e. when the RC was at least five times longer than the VP). The same was true for items belonging to the “other” predicate type: 67% of tokens with length ratios of 0.2 or less were

instances of RCE, but this immediately dropped to 29% for ratios between 0.2 and 0.4 (or 1:5 and 2:5) and less than 1% for items of length ratios greater than 0.6 (or 3:5). However, for tokens with an indefinite subject NP and a passive/presentative predicate, RCE was strongly preferred for items with length ratios up to 0.8, or 4:5 (Figure 6). This pattern suggests that grammatical weight plays the strongest role in predicting extraposition status for ratios less than 0.2, or 1:5 (where extraposition is usually preferred) and for ratios greater than 0.8, or 4:5 (where extraposition rarely ever occurs). For ratios between 0.2 and 0.8 (1:5 and 4:5), extraposition status appears to depend more on definiteness and predicate type than on grammatical weight.

These descriptive data suggest that length ratio, definiteness, and predicate type are important predictors of extraposition status, with possible interactions between definiteness and predicate type (Figure 2) and between definiteness and length ratio (Figures 5-6). Effects of subject accessibility (dispreference for RCE with given subjects) and predicate accessibility (preference for RCE with superset mention predicates) are also apparent (Figure 3).

Figure 5: Percentage of RCE for increasing ratios of VP length to RC length, definite tokens only

[Figure 5 here]

Figure 6: Percentage of RCE for increasing ratios of VP length to RC length, indefinite and passive/presentative tokens only

[Figure 6 here]

To test which factors were statistically significant, we used a binary logistic regression model with stepwise selection (cf. Diessel 2008), calculated using the PROC LOGISTIC function in SAS. The dependent variable was extraposition status (RCE = 1, non-RCE = 0). The

primary indicator chosen for determining the predictive power of the model and of each of the independent variables was the area under the Receiver Operating Characteristic (ROC) curve, where 0.5 indicates that the model predictions were at chance and 1.0 indicates that the model was 100% accurate (Zou et al 2007). To find the best measure of grammatical weight for predicting these data, alternative measures of grammatical weight were first tested one by one in a logistic regression model with no other independent variables: VP-to-RC length ratio (VP length divided by RC length), VP length alone, RC length alone, and length difference (RC length minus VP length). The area under the ROC curve was highest (at 0.91) for length ratio, and therefore length ratio was used as the measure for grammatical weight in the full model. The other independent variables included in the full model were as follows: definiteness (indefinite = 1, definite = 0), predicate type (passive/presentative = 1, other predicate type = 0), subject accessibility (given = 1, other = 0), predicate accessibility (superset mention = 1, other = 0), and all possible interactions. Note that “superset mention” was chosen as the relevant measure for predicate accessibility, because of the paucity of examples of discourse-given predicates (only two) in the corpus sample.

Statistics for the significant factors are summarized in Table 3, following Diessel (2008: 482). Four of the five independent variables were found to be significant predictors of extraposition status at an alpha level of $p < 0.05$: length ratio, predicate type, definiteness, and predicate accessibility.⁶ In addition, there were significant interactions between length ratio and definiteness and between predicate type and definiteness. These main effects and interactions are indicated by the Wald X^2 values and associated p-values in Table 3. Subject accessibility (givenness of the subject) was not a significant factor. The regression coefficients in Table 3 indicate whether RCE is more or less likely given a particular value of an independent variable.

The positive regression coefficients for definiteness, predicate type, and predicate accessibility show that indefinite subject NPs, passive or presentative predicates, and superset mention predicates each increase the likelihood of RCE. The negative value for length ratio (a continuous variable) shows that higher length ratios are associated with a *lower* likelihood for RCE (see Figure 4 above). As shown by the ROC curve values, length ratio was the strongest predictor of extraposition status, followed by predicate type, definiteness, and predicate accessibility. The overall prediction accuracy of the model, as indicated by the area under the ROC curve after inclusion of all the independent variables, was 0.9613 out of 1.0. Thus, the model was highly accurate for predicting extraposition status on this dataset.

Table 3: Statistically significant factors from logistic regression analysis (n = 345)

Independent variable	Wald X^2	p-value	Regression coefficient	Area under ROC curve when added first
Length ratio	$X^2 = 7.207$	p = 0.007	-2.562	0.918
Predicate type	$X^2 = 2.736$	p = 0.098	1.218	0.745
Definiteness	$X^2 = 8.246$	p = 0.004	3.062	0.711
Predicate accessibility	$X^2 = 5.012$	p = 0.025	1.869	0.579
Interaction: Predicate type * Definiteness	$X^2 = 6.276$	p = 0.012	3.529	0.748
Interaction: Length ratio * Definiteness	$X^2 = 7.030$	p = 0.008	-5.639	0.425

In summary, the statistically significant effects from the logistic regression analysis are consistent with our hypotheses, and also fill in details for which we had no clear predictions. Significant effects for length ratio, predicate type, definiteness, and predicate accessibility were as predicted in hypotheses (1-4) above. Subject accessibility (hypothesis 5) was not a statistically significant predictor of RCE, but numerically, the trend was in the expected direction: RCE occurred less often with given subject NPs than with new subject NPs. Interactions between predicate type and definiteness and between length ratio and definiteness, as illustrated in Figures 3, 5, and 6, were also found to be significant. In terms of relative strength, grammatical weight (length ratio) was the most reliable predictor of extraposition status, followed by predicate type, definiteness, and predicate accessibility.

X.3 Qualitative analysis of exceptional cases

The trends reported in the quantitative analysis were in the direction that previous studies of extraposition would lead us to expect. However, previous studies have relied on features or properties assumed to occur *consistently* in RCE tokens, and have not considered the independent effect of grammatical weight. If we see the formal and discourse-pragmatic features identified with RCE (discourse-new subject, presentative verb, heavy relative clause) as properties of an RCE *prototype*, what do we learn from examining ‘outliers’—exceptional cases in our corpus data?

Exceptional cases were defined as RCE items that lacked some of the theoretically significant features predicted by previous discourse-based analyses of RCE. We retrieved exceptional cases by using a SAS script to identify all RCE tokens that were incorrectly

predicted by the logistic regression model. Subsequently, all of the other RCE cases were examined manually to identify any additional exceptions. A total of 16 out of 53 RCE items were identified as “false negatives”—RCE tokens that were predicted by the model to have a non-RCE structure. An additional three items were then identified manually as exceptional, based on their apparent discourse function.

Of the 16 RCE items that the model failed to predict, at least six were theoretically unproblematic. In (7a-b), for example, all of the discourse and morphological features appropriate for RCE were present; e.g., both *changes* in (7a) and *some friends* in (7b) denote discourse-new entities. The model likely failed in these cases because of the relatively high VP-to-RC length ratio, which would have favored non-RCE structure, and which indeed makes these sentences sound rather awkward.

- (7) a. This is because changes were made to the standard rate contributions paid by employees that do not affect the married woman’s reduced rate contributions.
(written)
- b. We’ve got some friends coming to supper whose daughter’s there so I can question tomorrow so I can question her about it. (spoken)

More common among the exceptions were sentences that appeared to have topical subjects and focal predicates, as is more typical of non-RCE sentences. In each of the examples in (8a-c), the definite subject NP appears to denote a sentence topic, with focal stress falling somewhere on the predicate. Note, in addition, that (8a-b) exceptionally have copular predicates.

- (8) a. A: The one you did last time was for my wife actually, which was F name.
B: That probably explains why I can't find it under H. (spoken)

- b. As you can imagine the first few days will be a bit hectic, during which time I will be ringing you , and every client personally to invite you into my office.
(written)
- c. In aeolian environments the sand is blown until it accumulates, which can take on various features e.g. barchans, etc. (written)

The examples in (8) allow a non-restrictive interpretation of the relative clause, in which it expresses a distinct assertion from that of the matrix predicate. Accordingly, (8a-b) appear equivalent to (9a-b).

- (9) a. A: The one you did last time was for my wife actually. It was F name.
B: That probably explains why I can't find it under H.
- b. As you can imagine the first few days will be a bit hectic. During this time I will be ringing you, and every client personally to invite you into my office.

If the RCE construction allows only restrictive relative clauses, as is commonly assumed in the syntactic literature (see Baltin 2006), (8a-b) can perhaps be treated as something other than RCE. It is worth noting, however, that our corpus sample gave us no clear grounds for eliminating non-restrictive relative clauses: both (intuitively) restrictive and non-restrictive clauses modified a common noun and were introduced by a *wh* relative pronoun. Therefore, it appears problematic to exclude items like (8a-b) from analysis based on suppositions about the discourse function of relative clauses within RCE.

We will close with a final illustration of the challenges inherent in describing RCE function in (potentially) disfluent spoken English, where RCE may serve as a form of self-repair. In (10a), the subject NP, *the best singer*, is definite and has a discourse-active, topical referent,

and the predicate *is this Olaf Bergh*, appears to be focal. In addition, the relative clause is introduced by the complementizer *that*, preventing a non-restrictive reading, and the VP-to-RC length ratio of 4:3 is one that should disfavor RCE.

- (10). a. The best singer is this Olaf Bergh that I've seen. (spoken)
- b. ?The singer is this Olaf Bergh that I like best. (constructed example)

At least out of context, (10a) seems to us to be only marginally acceptable. What might be happening here? In (10a), the speaker seems to add the restrictive relative clause *that I've seen* as an afterthought, in order to qualify the claim made in the previous statement. As Wasow et al. 2011 observe, relative clauses are a common means of restricting the reference set over which a superlative applies, making a relative clause a natural choice of structure for such a qualifying statement. Consistent with this possibility, (10a) seems more felicitous than the constructed (10b), in which the subject contains no superlative and the relative clause is not used as a qualification. Thus, the sentence-final relative clause in (10a) appears to be a conventional self-repair, while that in (10b) is not. Support for this conjecture comes from a study of RCE comprehension by Levy et al. (2012), in which reading times were found to be faster for those RCE sentences in which the NP set up a strong expectation for a following relative clause (e.g., *only those executives...*), while RCE sentences were read more slowly than non-RCE sentences under weak-expectation conditions (e.g., *the executives...*).

In this section, we have discussed several cases of RCE from the corpus that appear to be problematic for current theories of RCE. Especially challenging are those tokens of RCE containing a topical subject and a focal predicate. Such cases call for an approach that

acknowledges both the effect of grammatical weight and the range of discourse functions extraposition may perform.

X.4 General discussion and conclusions

The current study reveals a complex interplay among several different factors contributing to speakers' and writers' choice of RCE as against non-RCE word order in English. Overall, there was a strong preference for RCE when the relative clause was at least five times longer than the VP (length ratio less than 0.2), and a strong preference for canonical order when the relative clause was the same length or shorter than the VP (length ratio 0.8 or higher). For those items with length ratios falling in the middle range (between 0.2 and 0.8), choice of structure appeared to depend primarily on the definiteness of the subject NP and on the type of predicate occurring in the main clause. Items with an indefinite subject NP and a passive or presentative main verb were much more likely to contain RCE than were items with other combinations of features. The accessibility of the predicate also had a small but significant effect: RCE was more likely with superset-mention predicates than with new (no prior mention) predicates. In short, it appears that length ratio sets soft limits on RCE based on ease of processing, while discourse-related factors regulate choice of structure within these limits.

More generally, this pattern of results appears to reflect a strategy by which speakers/writers resolved potential conflict between grammatical weight and discourse factors by giving preference to each under different conditions: grammatical weight was given priority in almost all cases, while discourse factors were given priority only for those tokens that fell within

a relatively narrow range of length ratios—those that might be considered neutral with respect to RCE. However, exceptional cases discussed in section X.3 show that this tradeoff was not always straightforward. For example, the sentences in (8a-b) show that occasionally, information-packaging considerations prevailed in licensing an RCE clause even when the length ratio favored a non-RCE structure. Future studies are needed to refine the conditions under which discourse factors may prevail over grammatical weight.

The results of the current study also have suggestive parallels with a corpus study by Jan Strunk on relative clause extraposition in German (Strunk, this volume). Similar to the current study, but on a larger scale (1300 tokens from the Tübingen Treebank of Written German), Strunk's study investigated a number of factors that independently contribute to writers' choice of extraposed or non-extraposed order in German. Unlike the current study, Strunk's study was not restricted to subject-modifying relative clauses but also included various kinds of complement-modifying relative clauses. Perhaps for this reason, in combination with language-specific differences, RCE was much more common in the German corpus. However, the results of the two studies are strikingly similar. As in the current study, factors related to grammatical weight (extraposition distance and relative clause length) were among the strongest predictors of extraposition. And as in the current study, discourse-related factors including definiteness and position of the relative clause antecedent (head noun) within the sentence were significant predictors of extraposition status, independent of grammatical weight. Like the English RCE option, the German RCE option was more likely to be chosen when the predication contained an indefinite antecedent than when it contained a definite antecedent. More generally, both studies suggest that speakers and writers are simultaneously sensitive to several different kinds of factors when making structural choices in language production, thus supporting the general approach of

several recent studies of word-order alternations (Arnold et al 2000; Bresnan and Ford 2010; Gries 2003; Lohse et al. 2004; Rosenbach 2005).

In addition to highlighting the interacting factors at play in the selection of an extraposed structure, the current study also calls into question common theoretical assumptions regarding the discourse function of RCE in English. Although a majority of tokens in the corpus were compatible with the predominant view of RCE as a presentational construction (Section X.2.2), a significant minority of RCE tokens (about 20%) appeared to have topical subjects and focal predicates (Section X.3). In contrast to previous analyses of RCE, which propose certain invariant use conditions (e.g., Rochemont and Culicover's claim that predicates of RCE sentences *must* be directly or indirectly c-construable), our results point toward a revised theoretical approach that allows for more flexibility in the way that extraposition is licensed. It is important to note that while length ratios favoring RCE (i.e., a short VP and long relative clause) can help explain a subset of these exceptional cases, other cases, including (8a) and (10a) above, cannot be explained in terms of a simple trade-off between discourse factors and weight. For (8a) and (10a), both the length ratios and the discourse properties should have favored a non-RCE structure. Thus, if only one discourse function were available for this construction, these cases would be difficult to accommodate. To capture the range of functions served by RCE, one could take a procedural approach, in which possible output forms are evaluated by means of ranked constraints, as in Optimality Theory (e.g. Bresnan et al 2007), or a declarative approach, in which constructions with shared formal, semantic or use conditions (e.g., various RCE constructions) are organized in a type hierarchy, as in classic Construction Grammar (e.g., Fillmore 1999), Head-Driven Phrase Structure Grammar (e.g., Malouf 2003), and Sign-Based Construction Grammar (e.g., Sag 2010). While a theoretical account is beyond the scope of this

short chapter, we tend to favor some version of the latter approach, in which non-prototypical construction types can inherit basic properties from the prototypical construction while lacking one or more of the properties that characterize the prototypical case. Cases like (8a-b) and (10a) could then be viewed as instantiations of a minor construction which inherits the syntactic properties of the basic RCE construction, while having a distinct discourse profile.

The need for a more flexible mapping between syntactic form and discourse function relates to a more general issue in research on non-canonical constituent order. Despite the recent trend toward multi-factorial studies of the kind reported here and in Strunk's chapter, many studies of non-canonical constituent order still assume a relatively narrow domain of explanation (e.g., syntax, semantics, and/or discourse information structure), based on a few selected examples. As we hope to have shown here, however, multi-factorial studies examining naturally occurring language have much to contribute to theories of how non-canonical constituent-order patterns are used and represented.

Acknowledgements

We gratefully acknowledge the help of our research assistants Maryana Bendus, Corinne Feight and Sunny Park, as well as our statistics consultants, Kenny Wakeland and Gayla Olbricht. We also thank colleagues and students who commented on earlier versions of this work, especially Jill Duffield, Roger Levy, Jack Hawkins, Brian MacWhinney, Edith Moravcsik, Jan Strunk, Tom Wasow and two anonymous reviewers. This research was jointly funded by the Department of English and the Linguistics Program at Purdue University and the Department of Linguistics at the University of Colorado Boulder.

References

- Arnold, Jennifer E., Thomas Wasow, Anthony Losongc, and Ryan Ginstrom (2000). 'Heaviness vs. newness: the effects of structural complexity and discourse status on constituent ordering', *Language* 76: 28-55.
- Baltin, Mark (1981). 'Strict bounding', In C. L. Baker and J. McCarthy (eds.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press, 257-295.
- Baltin, Mark (2006). 'Extraposition', In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax 2*. Malden, MA: Blackwell Publishing, 237-271.
- Bresnan, Joan, Ashwini Deo, and Devyani Sharma (2007). 'Typology in variation: a probabilistic approach to *be* and *n't* in the Survey of English Dialects', *English Language and Linguistics* 11: 301-346.
- Bresnan, Joan and Marilyn Ford (2010). 'Predicting syntax: processing dative constructions in American and Australian varieties of English', *Language* 86: 186-213.
- Carlson, Gregory (1977). *Reference to kinds in English*. PhD thesis, University of Massachusetts at Amherst.
- Culicover, Peter W. and Ray S. Jackendoff (2005). *Simpler syntax*. Oxford: Oxford University Press.
- Culicover, Peter W. and Michael S. Rochemont (1990). 'Extraposition and the complement principle', *Linguistic Inquiry* 21: 23-47.
- Diessel, Holger (2008). 'Iconicity of sequence: a corpus-based analysis of the positioning of temporal adverbial clauses in English', *Cognitive Linguistics* 19: 465-490.
- Francis, Elaine J. (2010). 'Grammatical weight and relative clause extraposition in English', *Cognitive Linguistics* 21: 35-74.

- Fillmore, Charles J. (1999). 'Inversion and constructional inheritance'. In G. Webelhuth, J.-P. Koenig and A. Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*. Stanford, CA: CSLI Publications, 113-128.
- Gibson, Edward (1998). 'Linguistic complexity: locality of syntactic dependencies', *Cognition* 68: 1-76.
- Gregory, Michelle L. and Laura A. Michaelis (2001). 'Topicalization and left dislocation: a functional opposition revisited', *Journal of Pragmatics* 33: 1665-1706.
- Gries, Stefan (2003). *Multifactorial analysis in corpus linguistics: A study of particle placement*. New York: Continuum Press.
- Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski (1993). 'Cognitive status and the form of referring expressions in discourse', *Language* 69: 274-307.
- Hawkins, John A. (1994). *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Huck, Geoffrey J. and Younghee Na, (1990). 'Extraposition and focus', *Language* 66: 51-77.
- Kay, Paul and Ivan A. Sag (2012). 'Cleaning up the big mess: discontinuous dependencies and complex determiners'. In H. Boas and I.A. Sag (eds.), *Sign-based construction grammar*. Stanford: CSLI Publications, 229-256
- Kayne, Richard S. (1994). *The antisymmetry of syntax*. Cambridge, MA: The MIT Press.
- Kuno, Susumu and Takami, Ken-ichi (2004). *Functional constraints in grammar: on the unergative-unaccusative Distinction*. Amsterdam: John Benjamins.
- Lamers, Monique J.A. and Helen de Hoop (this volume). 'Animate object fronting in Dutch: a

- production study’.
- Lambrecht, Knud (1994). *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Levy, Roger, Evelina Fedorenko, Mara Breen, and Edward Gibson (2012). ‘The processing of extraposed structures in English’. *Cognition* 122: 12-36.
- Lohse, Barbara L., John A. Hawkins, and Thomas Wasow (2004). ‘Domain minimization in English verb-particle constructions’, *Language* 80: 238–261.
- Malouf, Robert. (2003). ‘Cooperating constructions’. In E. J. Francis and L A. Michaelis (eds.), *Mismatch: form-function incongruity and the architecture of grammar*. Stanford, CA: CSLI Publications, 403-424.
- Michaelis, Laura A. and Hartwell S. Francis (2007). ‘Lexical subjects and the conflation strategy’. In N. Hedberg and R. Zacharski (eds.), *The grammar-pragmatics interface: essays in Honor of Jeanette K. Gundel*. Amsterdam: John Benjamins, 19-48.
- Nelson, Gerald, Sean Wallis, and Bas Aarts (2002). *Exploring natural language: the British component of the International Corpus of English*. Amsterdam: John Benjamins.
- Rochemont, Michael S. (1986). *Focus in generative grammar*. Amsterdam: John Benjamins.
- Rochemont, Michael S. and Culicover, Peter W. (1990). *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Rosenbach, Anette. (2005). ‘Animacy versus weight as determinants of grammatical variation in English’, *Language* 81: 613-644.
- Ross, John R. (1967). *Constraints on variables in syntax*. PhD thesis, Massachusetts Institute of Technology.
- Sag, Ivan A. (2010). ‘English filler-gap constructions’, *Language* 86: 486-545.

- Stallings, Lynne M. and Maryellen C. MacDonald, (2011). 'It's not just the "heavy NP": relative phrase length modulates the production of heavy-NP shift', *Journal of Psycholinguistic Research* 40: 177-187.
- Stallings, Lynne M., Maryellen C. MacDonald, and Pedraig G. O'Seaghdha, (1998). 'Phrasal ordering constraints in sentence production: phrase length and verb disposition in heavy-NP shift', *Journal of Memory and Language* 39: 392-417.
- Strunk, Jan (this volume). 'A statistical model of competing motivations affecting relative clause extraposition in German'.
- Takami, Ken-ichi (1999). 'A functional constraint on extraposition from NP'. In A. Kamio and K. Takami (eds.), *Function and structure: In honor of Susumu Kuno*. Amsterdam: John Benjamins, 23-56.
- Wasow, Thomas (2002). *Postverbal behavior*. Stanford: CSLI Publications.
- Wasow, Thomas and Jennifer Arnold (2003). 'Post-verbal constituent ordering in English'. In G. Rohdenburg and B. Mondorf (eds.), *Determinants of grammatical variation English*. The Hague: Mouton, 119-154.
- Wasow, Thomas, T. Florian Jaeger and David Orr (2011). 'Lexical variation in relativizer frequency'. In H.J. Simon, and H. Wiese (eds.), *Expecting the unexpected: exceptions in grammar*. Berlin: de Gruyter Mouton, 175-195.
- Zou, Kelly H., A. James O'Malley and, Laura Mauri (2007). 'Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models', *Circulation* 115: 654-657.

Figure 1: Definiteness, predicate type, and discourse status for RCE tokens (n = 53) and non-RCE tokens (n = 292)

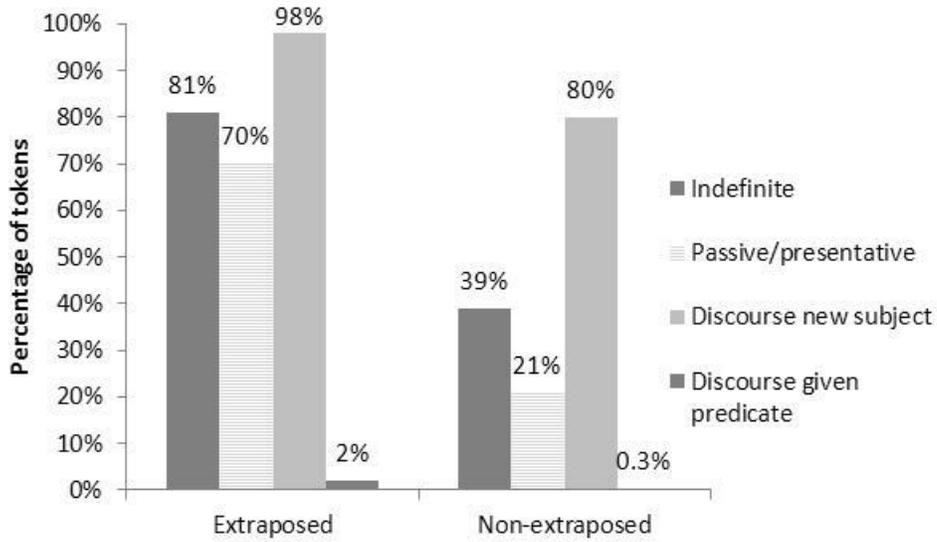


Figure 2: Percentage of RCE for four categories of definiteness and predicate type

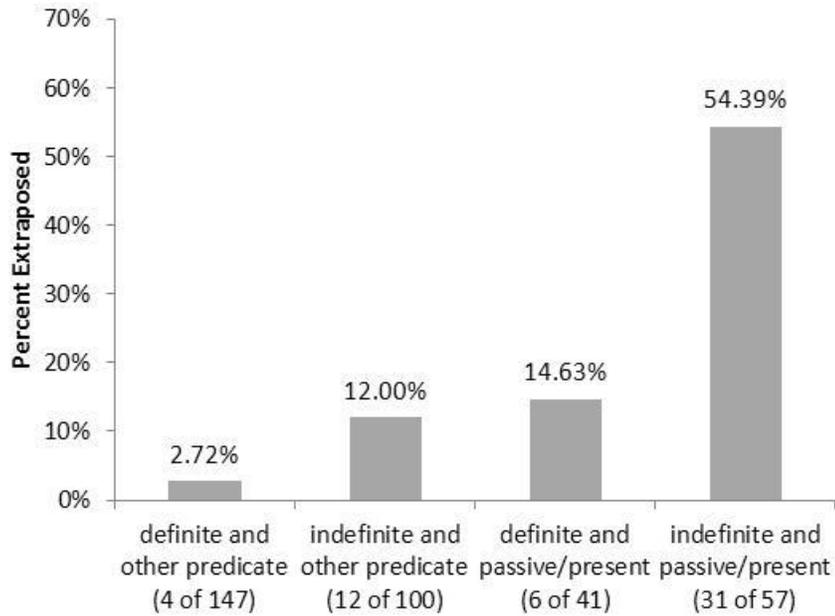


Figure 3: Percentage of RCE by discourse status of subject and predicate

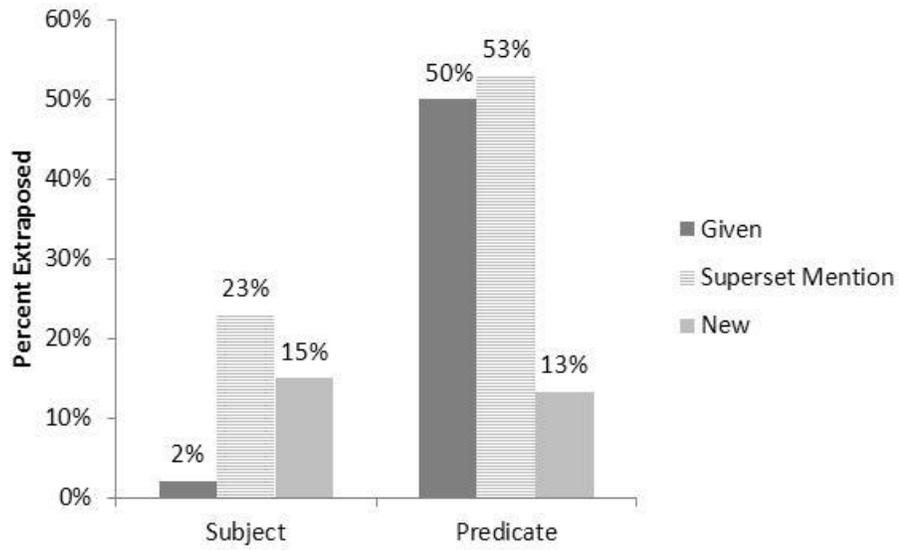


Figure 4: Percentage of RCE for increasing ratios of VP length to RC length

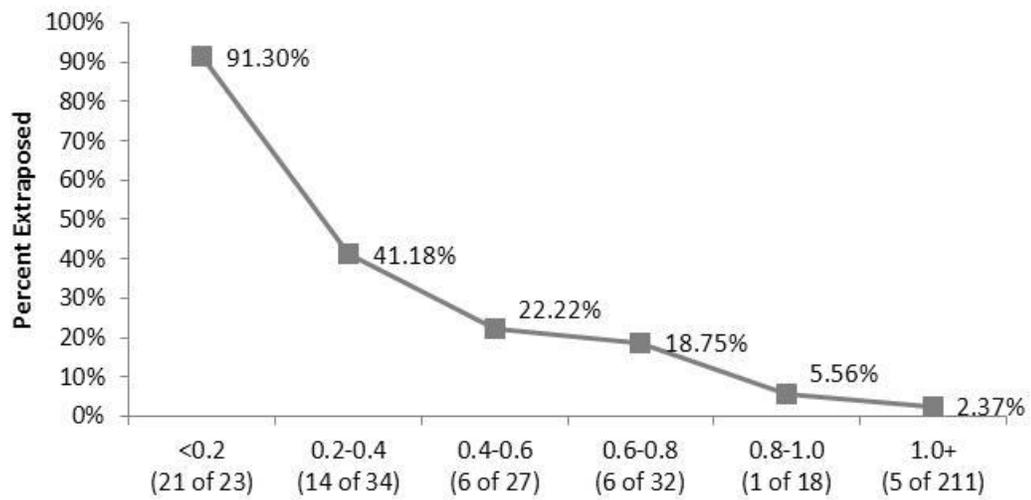


Figure 5: Percentage of RCE for increasing ratios of VP length to RC length, definite tokens only

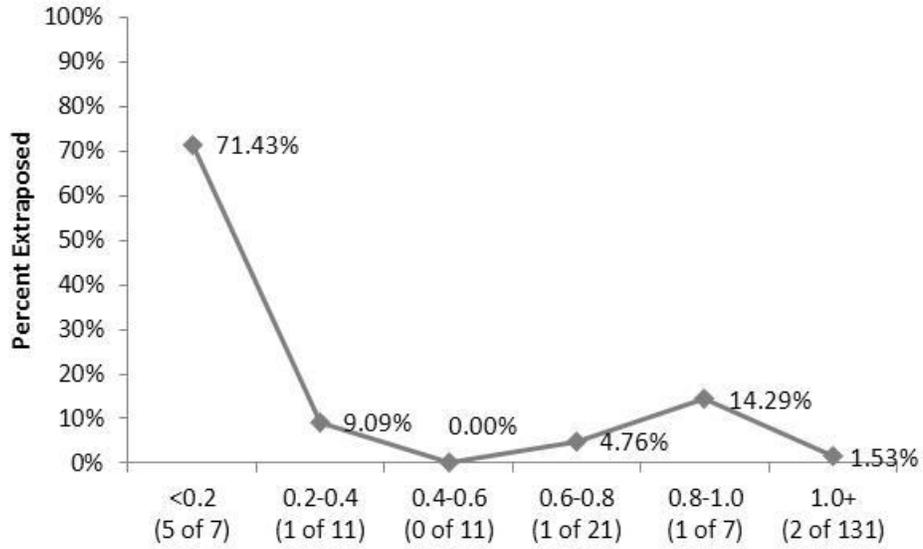
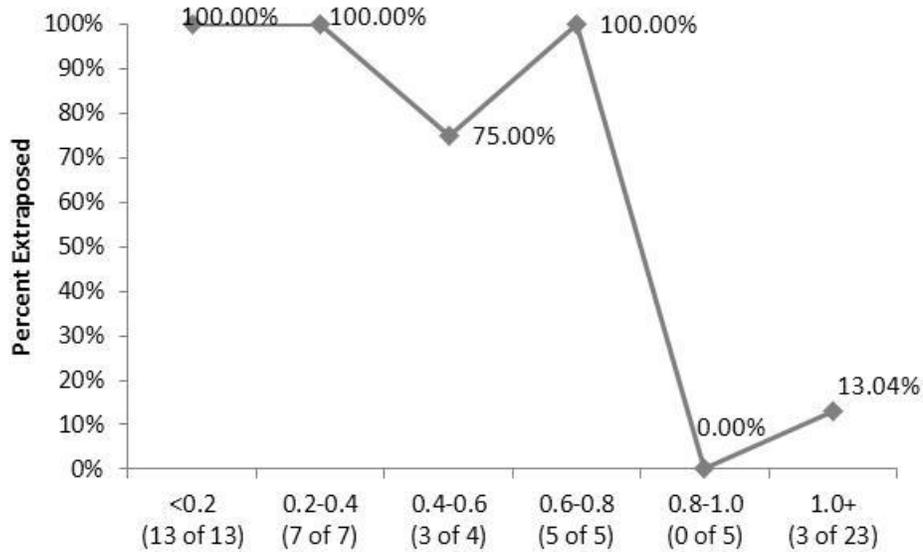


Figure 6: Percentage of RCE for increasing ratios of VP length to RC length, indefinite and passive/presentative tokens only



Footnotes

¹ The International Corpus of English Great Britain (Nelson, Wallis and Aarts 2002) includes about one million words of British English in a variety of genres of both speech and writing. All example sentences culled from the corpus will be indicated with the abbreviation ICE-GB in parentheses, as in (1a). This is also the corpus that we used for the empirical study described in section X.2.

² The latter result is somewhat at odds with the reading time results of Levy et al.(2012), in which RCE sentences were processed more slowly than non-RCE sentences in the absence of facilitating cues. However, the Levy et al. study had a different task (word-by-word reading) and did not include any weight manipulation.

³ For the purposes of statistical analysis, VP-to-RC length ratio is coded as a proportion. For example, for a ratio of 1:5, where the RC is five times longer than the VP, this was coded as 0.2.

⁴ In Figure 1, superset mention tokens are included in the discourse-new category. In Table 2, superset-mention tokens are counted separately.

⁵ One of the 292 non-RCE tokens was not analyzed for discourse status due to inadequate context.

⁶ As shown in Table 3, the p-value for predicate type of 0.098 was not significant due to the inclusion of the interaction between predicate type and definiteness in the model. When interactions are excluded, predicate type becomes highly significant at $p < 0.0001$.