The ellipsis alternation: remnants with and without prepositions

Joanna Nykiel

Abstract

The ellipsis alternation refers here to the previously unexplored variation between two kinds of ellipsis remnants whose correlates are PPs. One kind of remnant includes the preposition hosted by its correlate and the other does not. I explore determinants of this variation in English sluicing and Bare Argument Ellipsis, and show that they can be captured by three performance preferences whose presence has been independently observed in research on anaphora, constituent ordering, and priming (structural persistence), and whose motivation has been found in the efficiency of language processing. These preferences include reduction of the morphosyntactic and semantic content of an anaphor which has an accessible antecedent, adjacency of constituents which depend semantically on each other, and replication of the form of the correlate. I demonstrate that, together, these preferences can account for the gradience of the ellipsis alternation in corpus and experimental data. I further suggest that the strength of these preferences in various languages forms a basis for evaluating differences and similarities among them with respect to the availability of the ellipsis alternation. I conclude that the current data speak in favor of usage-based theories of grammar.

1 Introduction

There is growing recognition in the linguistic literature that speakers’ knowledge of language involves both categorical and non-categorical information. It is uncontroversial that categorical information framed in terms of categorical judgments of grammaticality is part of speakers’ knowledge, but the possibility that non-categorical or gradient information is part of that knowledge has also recently been supported by studies of syntactic variation (Arnold et al. 2000, Hawkins 2000, Wasow 2002, Bresnan 2007, Bresnan et al. 2007, Ford and Bresnan 2010, Rosenbach 2002, 2003, Szmrecsanyi 2005, 2006, Nykiel 2013a, Wolk et al. 2013). All of these studies show that speakers select one of two available alternatives, based on a variety of contextual factors, and some further demonstrate that speakers’ choices may be accurately predicted by statistical models of the data (Bresnan et al. 2007, Ford and Bresnan 2010). An interesting finding revealed by, for example, Bresnan (2007) is that constructions that are commonly assumed to be ungrammatical are in fact used and judged acceptable by speakers under certain conditions. Typically, it is the low frequency of such constructions or insufficient context
that triggers judgments of ungrammaticality, indicating that speakers possess gradient knowledge of language which is best captured by a probabilistic grammar (Manning 2003).

In this paper, I focus on what I dub the ellipsis alternation, syntactic variation seen in two elliptical constructions. One of them is sluicing, a construction where a wh-phrase is left stranded and has an overt correlate in the antecedent clause (Ross 1969). The meaning of the wh-phrase (which I will henceforth refer to as a remnant) is that of a wh-question. Two syntactic alternatives, a stranded wh-phrase with or without a preposition, are available in the context shown in (1) and (2), where the remnants’ correlates are PPs, with someone and on my other friends (here and elsewhere, correlates are indicated by italics). In the rest of this paper, I will refer to remnants with prepositions as PP remnants and to the other alternative as NP remnants.

(1) Katie is staying in Paris with someone, but I don’t know with who/who.

(2) A: I’m counting on my other friends.
   B: On what other friends?/What other friends?

The second construction is Bare Argument Ellipsis (henceforth BAE), illustrated in (3) and (4). Here, the remnant is an XP which answers the question posed in the antecedent clause (3)) or is a comment on the antecedent ((4)). As with sluicing, the correlate is a PP, although its constituents may be nonadjacent due to preposition stranding in the antecedent clause, as is the case in (3). Two alternatives are again available as remnants.

(3) A: What meds are you on?
   B: On none/None.

(4) A: Kim’s sister is proud of her for her modeling career.
   B: But not for telling the truth/telling the truth.

These two constructions receive a uniform syntactic analysis in Culicover and Jackendoff (2005), leaving little reason to doubt that they show similar behavior with respect to the variation in question. I keep them separate throughout except for the statistical analysis, as described in section 3.2.

No empirical studies of this variation in English are available, although it is unclear whether it is subject to categorical or "soft" grammatical constraints. It has been a standard assumption since Merchant (2001) that both NP and PP remnants are grammatical in English, because preposition stranding and pied-piping are available in corresponding nonelliptical clauses. In formal terms, this means that NP and PP remnants may be interpreted as deriving from nonelliptical clauses through movement and deletion and with or without preposition stranding, as shown in (5) and (6) (strikethrough indicates the deleted material).

(5) Katie is staying in Paris with someone, but I don’t know with who she is staying.

(6) Katie is staying in Paris with someone, but I don’t know who she is staying

---

1See also Merchant (2004) for an argument that BAE should be assimilated to sluicing. Analysis of BAE, however, is limited to the question and answer type shown in example (3).

2But see Nykiel (2013a) for Polish and Merchant et al. (2013) for German.
Merchant’s work does not go as far as to try to identify constraints on the appearance of NP and PP remnants, but one might expect to see parallels between constraints on preposition placement in nonelliptical clauses and constraints on the realization of ellipsis remnants. Such parallels are clearly visible in data such as (7) and (8). For these idiomatic combinations of verbs and prepositions, preposition stranding is mandatory, and NP remnants are much more natural than PP remnants, although the grammar of English allows both kinds of remnants here.

(7) A: *What did Ben fall for?* (cf. *For what did Ben fall?*)
    B: The very same trick as before. (cf. *For the very same trick as before*).

    b. Ben came *across a bunch of old documents*, but I don’t know whose. (cf. *Ben came across a bunch of old documents, but I don’t know across whose.*

However, there is also evidence that the variation seen in ellipsis is freer than the preposition placement possibilities in corresponding nonelliptical clauses. For instance, Chung et al. (1995) and Fortin (2007) point out that NP remnants may appear where preposition stranding is unavailable. This is illustrated in (9) and (10) from Chung et al. (1995: 273); interrogative clauses are shown together with the corresponding ellipses.

(9) a. *What circumstances will we use force under?*
    b. We will use force *under certain circumstances*, but we can’t say what.

(10) a. *What sense is this theory right in?*
    b. This theory is right *in some sense*, but I’m not sure what.

If we move beyond remnants with possible sentential sources, the question of the availability of preposition stranding in corresponding nonelliptical clauses does not arise. Consider (11) and (12), where the remnants following the phrases *let alone* and *but* have been argued to be nonsentential (Culicover and Jackendoff 2005, Toosarvandani 2013). And yet, NP remnants are no less natural here than they are in examples (7) and (8). Thus whatever mechanism is responsible for the degradation of PP remnants in (7) and (8) must be the same as the mechanism responsible for the degradation of PP remnants in (11) and (12).

(11) Ben will fall *for any trick*, let alone an old one/?for an old one.

(12) Ben didn’t come *across old documents* but toys/?across toys.

Drawing on experimental and corpus evidence, I demonstrate that the knowledge of when to use appropriate ellipsis remnants represents not categorical, but gradient information about the contexts that favor a particular type of remnant. This gradient information is motivated by three constraints on sentence processing (performance preferences) independently known to affect anaphoric constructions, constituent ordering, and language production. I further suggest that these constraints underlie differences among languages with respect to how freely both NP and PP remnants are available,
and conclude that the current data support usage-based theories of grammar (Manning 2003, Bod 2006, Bresnan and Hay 2008). The argument is that the constraints on the ellipsis alternation discovered here advance our understanding of the relationship between this alternation and preposition placement in nonelliptical clauses. Although this relationship might appear to be syntactic, it is not. Thus the grammar of English (and other languages) is arguably shaped by patterns of preference found in performance.

The rest of the paper is structured as follows. The next section describes data collected for this study and the statistical analysis of it. Section 3 details factors that I considered and tested as predictors of the variation between NP and PP remnants. Section 4 presents a mixed-effects model of the corpus data. In section 5, I compare, through regression modeling, speaker ratings for selected items from the corpus data with the corpus model’s probabilities for these items. In section 6, I examine the nature of the constraints on the variation between NP and PP remnants, and consider their implications for the typology of languages with respect to the realization of ellipsis remnants, theories of ellipsis, and theories of grammar. Section 7 concludes.

2 Data

The data I collected for this study come from three corpora of spoken American English: the Switchboard corpus (henceforth S), Santa Barbara (henceforth SB) and the Corpus of Contemporary American English (henceforth COCA). I first extracted all wh-phrases from the first two corpora and then identified those that were instances of sluicing with PP correlates. Next, I identified those wh-phrases which were embedded in interrogative clauses as prepositional objects and selected those for which BAE remnants served as responses. These data constitute 60% of the entire dataset. As for the third corpus, I extracted an equal sample of bare wh-phrases and which/what/whose-NP phrases from the spoken part of the corpus to avoid a potential bias toward a particular kind of wh-phrase. Given that COCA is not a parsed corpus, extracting all wh-phrases would have been quite a time-consuming process. YES/NO questions and elliptical responses to them are not part of the dataset.

The constructions I extracted from the corpora include ordinary remnants for which plausible sentential sources may be posited (see examples (5) and (6)) as well as remnants that lack obvious sentential sources (see examples (11) and (12)). While constructions of the latter type do not always receive the same amount of attention as constructions of the former type, there is no apparent reason to leave them out of the current analysis, particularly because, as suggested by examples (11) and (12), constraints on the use of NP and PP remnants generalize to all relevant constructions. I extracted the total of 411 ellipsis remnants, of which 276 (67.2%) were NP remnants, and coded them for factors to be catalogued in the next section. This sample provides enough combinations of factors expected to affect the realization of remnants to reveal the role each plays in the ellipsis alternation.

The statistical analysis of the data is a mixed-effects regression model predicting the realization of remnants. This procedure involves selection of the best-fitting model that can accurately predict the appearance of one of two alternatives based on a number of factors. Regression modeling (Baayen 2008) is gaining in popularity among experimental and corpus linguists investigating both synchronic and diachronic data (van Kemenade et al. 2007, Bresnan et al. 2007, Ford and Bresnan 2010, Wolk et
al. 2013). It is particularly useful as a means to explore instances of variation where several different factors are potentially involved and where statements as precise as possible are needed about which of these factors play a role and which do not. Because the variation between remnants I am exploring here is genuinely unexplored, it is important to provide reliable evidence for the reality of constraints that might be influencing it. I next turn to the factors included in the best-fitting model I developed for the data.

### 3 Factors included in the model

This section details the factors that are relevant to predicting how ellipsis remnants are realized. These include correlate informativity, construction type, structural persistence, semantic dependencies between lexical categories, and presence of ellipsis in the antecedent clause. They are factors included as fixed effects in the mixed-effects model of the corpus data, which I present later in this paper. For each of the factors, I provide my motivation for including it in the model and a brief description of the annotation process, as required.³

#### 3.1 Correlate informativity

There is evidence that the inherent accessibility of an entity influences language production where speakers have several syntactic options to choose from. Inherent accessibility is taken to be part of conceptual accessibility, which refers to "the ease with which the mental representation of some potential referent can be activated in or retrieved from memory" (Bock and Warren 1985: 50). Inherent accessibility is independent of the context in which an entity appears, but is determined by the entity’s semantic features, such as animacy, number, concreteness, and referentiality (Prat-Sala and Branigan 2000, Jaeger and Wasow 2008).⁴ Entities that are animate, singular, physical, or referential are inherently more accessible than entities that are inanimate, plural, abstract, or impersonal, and are therefore easier to retrieve. Inherent accessibility is known to affect word order variation, because inherently accessible entities tend to be assigned early sentence positions, serving as subjects (Prat-Sala and Branigan 2000). It is also known to play a role in relativizer omission in non-subject-extracted relative clauses, such that inherently more accessible subjects of relative clauses show low relativizer frequency (Jaeger and Wasow 2008). While the focus of this research is on differences in how the features contributing to inherent accessibility are specified, I focus here on the fact that some expressions are specified for these features, but others are underspecified for them. For instance, the correlates in examples (13)–(15) differ in terms of which features they are specified for.⁵

³In addition to the five fixed effects detailed below, I explored the possibility that the length of the remnant measured in words influences the pattern of preposition omission. I hypothesized that the longer the remnant the less likelihood of it being a PP. There was no empirical support for this effect, however.

⁴Prat-Sala and Branigan (2000) separate inherent accessibility out from derived accessibility, which is subject to contextual manipulations, while treating both as components of conceptual accessibility.

⁵Examples that come from any of the corpora used for the current analysis are annotated with an appropriate corpus source; if there is no annotation, the examples are constructed.
In (13), the correlate points to a set of viruses, and in (15), the correlate refers to the US states. But in (14), the correlate merely denotes a human, and hence animate, entity. The correlates in (13) and (15) are specified for more semantic features (e.g., concreteness in addition to animacy) than the correlate in (14).

These examples raise the question of whether the number of features a correlate is specified for influences its inherent accessibility. It is reasonable to assume that it does. This assumption fits well with research on memory retrieval showing that semantically and syntactically complex expressions receive stronger mental representations than expressions that are less complex, which aids in their future retrieval from memory (Craik and Lockhart 1972, Gallo et al. 2008, Hofmeister et al. 2007, Hofmeister et al. 2013, Hofmeister 2007, 2008, 2011). Specifically, it has been shown that more complex expressions have a processing advantage over less complex ones in filler-gap constructions, such as in (16). The more syntactically and semantically complex filler is the NP \textit{an alleged Venezuelan communist} and its less complex counterpart is the NP \textit{a communist}. When these NPs are retrieved at the gap site (following the verb \textit{banned}), the more complex one is read faster, an indication that it is easier to retrieve.

\begin{equation}
\text{It was a communist/an alleged Venezuelan communist who the members of the club banned from ever entering the premises. (Hofmeister 2011: 385)}
\end{equation}

Hofmeister (2011) demonstrates that such reading time differences persist even if fillers differ in semantic but not syntactic complexity. This is the case in (17). Note that the less complex of the two fillers, \textit{which person}, is specified for fewer semantic features than the other, because it is not specified for the type of individual. Hofmeister (2011) proposes that the greater number of features a phrase carries, the more unique its mental representation becomes, ensuring that it remains easily identifiable among the other stored representations in the discourse model.

\begin{equation}
\text{The lieutenant could not remember which person/which soldier the commander that was deeply respected ordered to scout the area ahead. (Hofmeister 2011: 388)}
\end{equation}

Given this research, I assume that remnants' correlates differ in inherent accessibility, depending on the set of features they are specified for.

Another question raised by the examples in (13)–(15) is whether ellipsis alternation is affected by differences in the number of features that correlates are specified for. I hypothesize that correlates with more features exhibit higher frequencies of NP remnants, because such correlates are more accessible and hence require fewer retrieval cues at the ellipsis site. Support for this hypothesis comes from two strands of research: processing of elliptical constructions (Martin and MacElree 2011) and nominal
anaphora (Ariel 1990, 2001). Martin and MacElree (2011) argue that antecedents for sluicing remnants are retrieved directly and not by a sequential search through the stored representations. What plays a role in such a direct retrieval process are both the antecedent’s mental representation and the retrieval cues provided by the remnant. By hypothesis, the quality of the antecedent’s mental representation will vary with the features borne by the correlate: the more features, the stronger the representation. As for the remnant, I assume that a PP remnant provides more retrieval cues than an NP remnant, given that it matches the syntactic category of its correlate, which is also a PP. Hence, a PP remnant should be found with less accessible antecedents more frequently than an NP remnant. Ariel’s (1990) work on nominal anaphora suggests that this prediction is correct.

Ariel (1990: 73) proposes a scale of nominal anaphoric expressions, ranging from the most explicit in morphosyntactic terms (a full name + modifier) to the least explicit (zero). The less explicit type of anaphoric expression serves to signal high-accessibility antecedents, whose retrieval is not costly; the more explicit type of expression signals low-accessibility antecedents, indicating a processing cost involved in reactivating them. Since, similar to nominal anaphors, the function of ellipsis remnants is to reactivate their antecedents, prepositions may be exploited as a means to manipulate the amount of retrieval cues. The presence/absence of prepositions in remnants may be couched in terms of lengthening or reducing the remnant. Such manipulations are also known from nominal anaphora. The pronunciation of an anaphor such as a pronoun can be lengthened if its antecedent occupies a nonsubject position (Brennan 1995). This is done in order to signal to the hearer that the antecedent is less accessible than perhaps expected, given the choice of the pronominal anaphor. Conversely, the pronunciation of explicit anaphors (e.g., proper names) is subject to reduction if the antecedent is perceived as highly accessible (Fowler et al. 1997). Beyond pronunciation, such adjustments persist at the morphosyntactic level. Proper names, for instance, refer to antecedents of low accessibility, according to Ariel (1990). However, there is a finer distinction to be made here: full names mark antecedents of lower accessibility than partial names (Mulkern 1996). Given that proper names have a range of possible realizations, using a first name instead of a full name produces a reduction in the morphosyntactic complexity of the name. Assuming, with Payne and Huddleston (2002: 518), that proper names are NPs, first names consist of simple heads, while full names represent ”composite head structures”. The choice between a first and a full name thus has morphosyntactic and semantic consequences, that is, using the first name reduces the structure of this NP to just a head, but using the full name expands the NP’s structure to a composite head.

In the rest of this paper, I use the term informativity rather than complexity when referring to the number of features that correlates are specified for, which avoids potential confusion with syntactic complexity. I formalize the notion of informativity as a predictor of which correlates are likely to appear with which remnants. Let us assume first that the number of syntactic and semantic features that a correlate may be specified for is maximally ten. This is an arbitrary number, given that in principle a large number of features could be available, provided that the correlate can still give rise to a Question under Discussion (QUD). These features are listed in Table 1.

---

6QUDs are sets of alternatives that can serve as answers to a question; they can be introduced directly by a question or indirectly by grammatical devices, such as indefinite reference (Clifton and Frazier 2012).
Some of these features require immediate clarification. To capture the implicational relationship between the features humanness and animacy (nonhuman is not necessarily animate, but human is always animate), I analyze them as separate features. To see this, consider the interrogative pronouns *who* and *what*. While the former is specified for both humanness and animacy, the latter only for humanness. Further, I analyze all nonhuman entities as specified for natural gender (neuter). For human entities, natural gender is specified only if the sex is specified. For example, the interrogative pronoun *who* is analyzed as not specified for natural gender. For the attributive feature, I assume that only one adjective from this set is present in a correlate at a time. Note also that some features are borne by all correlates (CAT, number, referent, humanness), some are borne by some correlates (animacy, concreteness, natural gender, case, attributive), and one feature is borne by no English correlate (grammatical gender). Languages differ in the kind of features which phrases such as an NP or indefinite pronoun may be specified for. For purposes of this discussion, I include the ten features listed above to ensure generality beyond English. As will become clear shortly, it is not the maximal number of features available in a language that matters but the number of features actually borne by one phrase relative to the number of features actually borne by another phrase in a single language.

This set of features employs both syntactic and semantic features to distinguish among various correlates. However, it could be that the overall informativity of a correlate is reducible to its semantic features, and hence only those might ultimately be required for measuring informativity (recall the semantic complexity effects found in filler-gap constructions). Since we do not yet understand whether it is the semantic or syntactic features of a correlate that modulate its informativity and interact with retrieval, especially in contexts where several candidates for retrieval are present in the antecedent at the same time, triggering interference effects, I chose to include both semantic and syntactic features already in the current analysis, providing a basis for future research.

The semantic features I have chosen for the analysis are fairly general, used in componential analysis, and include features known to contribute to inherent accessibility. My purpose is to establish whether presence or absence of these features is enough

---

<table>
<thead>
<tr>
<th>Feature</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic category (CAT)</td>
<td>NP, indefinite pronoun, which-NP, or bare wh-phrase</td>
</tr>
<tr>
<td>Grammatical gender</td>
<td>masculine, feminine, or neuter</td>
</tr>
<tr>
<td>Number</td>
<td>singular or plural</td>
</tr>
<tr>
<td>Case</td>
<td>nominative, genitive, etc.</td>
</tr>
<tr>
<td>Animacy</td>
<td>animate or inanimate</td>
</tr>
<tr>
<td>Concreteness</td>
<td>physical or abstract</td>
</tr>
<tr>
<td>Humanness</td>
<td>human or nonhuman</td>
</tr>
<tr>
<td>Natural gender</td>
<td>masculine, feminine, or neuter</td>
</tr>
<tr>
<td>Attributive</td>
<td>size, shape, color, age, etc.</td>
</tr>
<tr>
<td>Referent</td>
<td>individual, singleton or nonsingleton set</td>
</tr>
</tbody>
</table>

Table 1: Features and their levels

In the case of BAE, QUDs are introduced directly by the antecedent clauses, but in the case of sluicing, they are introduced via indefinite phrases functioning as correlates.
to affect the distribution of remnants above the chance level. More concretely, I ask whether an increase in the subset of the current features which is borne by correlates causes a corresponding increase in the frequency of NP remnants. We could add more specific features to the set, such as the category or type of object or individual: a hammer, for instance, would be a type of tool in addition to being inanimate, physical, nonhuman, and neuter. If such more specific features were included, they would lead to an increase in the informativity of correlates hosting NPs compared to those hosting indefinite pronouns. If informativity effects are observed with the current set of features, addition of more specific ones would amplify these effects. An extension of the current study would be to explore how the different levels, and not just the specification vs. nonspecification for particular features, affect the ellipsis alternation. One question to ask in future research is, for instance, does a correlate denoting an animate, singular or physical entity exhibit a higher frequency of NP remnants than a correlate denoting an inanimate, plural or abstract entity? The informativity of many phrases will be identical in the current system so long as they are specified for the same number of features, whether these features differ or not and whether the levels are the same or not, if the features do not differ. Thus, which woman and which famous senator do not differ in informativity (although only the former is specified for natural gender and only the latter contains an attributive term), and nor do which woman and which man.

The correlate’s informativity increases in proportion to the number of features that it is specified for. For each phrase serving as a correlate, we can calculate its informativity score by dividing the number of features it is actually specified for by the maximal number of features.\footnote{I thank Philip Hofmeister (personal communication) for suggesting a solution along these lines.} Let us use the examples in (13)–(15) to illustrate this. The correlate in (13) encodes 8 features out of 10: CAT, number, humanness, animacy, concreteness, natural gender, an attributive term, and a referent.\footnote{No English NP encodes grammatical gender or case (except for genitive phrases). Thus the maximum informativity score for an English NP is 0.80} Its informativity score is 0.80. The correlate hosting an indefinite pronoun in (14) encodes 5 features out of 10: CAT, number, animacy, humanness, and a referent. Its informativity score is 0.50. Finally, the correlate with a what-NP phrase in it like that in (15) encodes 7 features out of 10 (CAT, number, humanness, animacy, concreteness, natural gender, and a referent), receiving the informativity score of 0.70. I coded the corpus data for correlate informativity by assigning to each item an appropriate score. I assumed, following Ariel’s accessibility scale, that proper names, whether or not they included any attributive terms, had the maximum informativity score of 0.80 available in English, and so did first-person personal pronouns (me or us) referring to the speaker in a given discourse situation. With this system in place, we can proceed to explore the relationship between a correlate and remnant.\footnote{Notice that the proposed system leaves the informativity of the remnant out of the analysis. Which-NP phrases are more informative than bare wh-phrases. To see why, consider that the informativity score of an English which-NP phrase is at least 0.70 (the same as that of an unmodified NP), and the informativity score of a bare wh-phrase is 0.40 or 0.50 (the same as that of an indefinite pronoun). This difference in informativity is linked to the type of correlate, that is, correlates that host NPs strongly prefer which-NP phrases as remnants, but correlates that host indefinite pronouns strongly prefer bare wh-phrases as remnants (Nykiel 2013b). Thus neither type of wh-remnant can be analyzed as inherently biased toward preposition omission or retention outside of its relationship with the correlate. The clear intuition that emerges from this}
It has not been proposed yet that correlate informativity plays a role in the realization of ellipsis remnants in a language like English. My goal here is to test whether it does by treating correlate informativity as a distinct factor. The raw frequencies for the full dataset arranged according to correlate informativity scores appear in Table 2.¹⁰ I also coded these data for the number of words per correlate (with prepositions excluded) as a way of exploring weight effects. No such effects were observed in the data, and for this reason, I eliminated weight from the final model, leaving only informativity as a numeric factor.

<table>
<thead>
<tr>
<th>Informativity score</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>77(36.3%)</td>
<td>135(63.7%)</td>
<td>212(100%)</td>
</tr>
<tr>
<td>0.70</td>
<td>26(26%)</td>
<td>74(74%)</td>
<td>100(100%)</td>
</tr>
<tr>
<td>0.80</td>
<td>32(32.3%)</td>
<td>67(67.7%)</td>
<td>99(100%)</td>
</tr>
</tbody>
</table>

Table 2: Realization of remnants by correlate informativity score

Note that, as expected, an increase in the frequency of NP remnants correlates with increasing informativity scores. Regression modeling of these data will evaluate whether this correlation is statistically significant.

3.2 Construction type

Beyond the basic construction types of sluicing and BAE, I identified split questions and reprise questions. I use the term construction type pre-theoretically here, as a means to capture different functions of the four categories, as detailed below.

Split questions are constructions rarely discussed in the ellipsis literature (see Arregi 2010). Consider (18).

(18) This cake is filled with what, whipped cream?

They appear to be similar to BAE. In BAE, remnants are analyzed as separate utterances from the antecedents, while in split questions, they are either analyzed as separate clauses (Arregi 2010) or parts of the antecedent clauses they are adjoined to (Camacho 2002). The details of these analyses need not concern us here, but if we assume that remnants are not part of the antecedent clauses in split questions, then split questions resemble BAE. While both Arregi (2010) and Camacho (2002) assume that split questions require the presence of a comma separating the interrogative from the phrase that follows it, this seems to be a matter of punctuation. (18) can be modified such that it contains two separate interrogative clauses of which the second is the tentative answer to the first:

(19) This cake is filled with what? Whipped cream?

Compared to BAE, the salient feature of the types shown in (18) and (19) is that they occur within a single speaker’s turn, but they are difficult to tease apart based on their relationship is that which-NP phrases support reduction of the remnant through omission of prepositions better than bare wh-phrases primarily because they have informative correlates.

¹⁰Two correlates received the informativity score of 0.50, but I collapsed them with the 0.40 level in the table due to their low frequency.
intonational features; punctuating them according to the pattern in (18) or (19) is thus an arbitrary decision. I collapse both types as a single category of split questions in this paper. Split questions have a relevance to the current discussion because, like ordinary BAE and sluicing, they allow NP and PP remnants if PPs appear as correlates within the antecedent clauses.

Reprise questions deserve a more detailed treatment here. Ginzburg and Sag (2000) employ the term reprise questions (from Bolinger 1957) to refer to wh-phrases in situ which ask for clarification of some part of the prior utterance. The need for clarification results from either mishearing part of the utterance (echo questions) or failure to identify its referent (reference questions). For both kinds of reprise questions, elements of the prior utterance are repeated and the wh-phrases accented, while the intonation patterns differ. Echo questions have a "focus-associated rise with spreading high tone" and reference questions a "focus-associated fall with spreading low tone" (Ginzburg and Sag 2000: 256). (20) is an echo question, and (21) a reference question (emphasis original).

(20) A: Bustamante y Bacigalupo plays the violin badly.
    B: WHO plays the violin badly? (Ginzburg and Sag 2000: 255)

(21) A: They’re mad at Bustamente y Bacigalupo.
    B: WHO’s mad at Bustamente y Bacigalupo? (Ginzburg and Sag 2000: 255)

In a corpus of transcribed conversation, that is, without access to intonation patterns, it may be difficult to separate echo questions out from reference questions, as in, for example, (23). In other cases, echo (22) and reference questions (24) are quite easy to identify in context.

(22) A: There is someone that I have just discovered like in the last couple of years and have you ever heard of Yani?
    B: Of who?
    A: Yani.
    B: The name sounds familiar. (S)

(23) A: But he couldn’t get along with Aguilar.
    B: With who?
    A: I mean Aguirre.
    B: Oh, Aguirre. (S)

(24) A: I think there are many Chinese women with that.
    B: With what? Tell me.
    A: With many abortions. (COCA)

As the examples in (22)-(24) illustrate, reprise questions can be elliptical (the kind relevant to the current study) and even look like sluicing if they contain wh-phrases. But they needn’t contain such phrases.\textsuperscript{11}

(25) A: And, okay, let’s see. Well, who do you like from that category?

\textsuperscript{11}Ginzburg (2012) refers to reprise questions like those in (22)-(24) as reprise sluices, and to those in (25)-(26) as reprise fragments.
B: From New Age?
A: Yeah. (S)

(26) A: And what do you think you’ll do with that?
B: With those degrees? (S)

Based on the largely parallel behavior of echo and reference questions with respect to the preference for PP remnants, I collapsed them into a single type of reprise questions. This type is distinguished here by its clarificational character and kept separate from BAE and sluicing, which only include nonclarificational utterances.

However, it is important to note that within the type of reprise questions, correlates may and often do vary in terms of informativity. In coding correlates for informativity, I assigned to all correlates for echo questions the lowest informativity score found in English (0.4), even if they contained proper names, as in (23). This assumption follows from the fact that the correlate in (23) has been misheard and hence its referent remains unidentifiable despite a high informativity score that a proper noun would otherwise receive. Unlike echo questions, reference questions feature correlates which have been heard correctly and whose referents may be more or less accessible as the consequence of the correlates’ informativity, although they have not yet been fully identified. Compare the correlate with that in (24) with the more informative correlates in your condition and on our list in (27)–(28). Therefore, I employed the same scale of informativity in the case of reference questions that I used for other constructions.

(27) A: You shouldn’t be playing in your condition.
B: What condition? (COCA)

(28) A: I’m putting you on our list.
B: What list? (COCA)

Isolating the different construction types makes it possible to assess the contribution of each to predicting speakers’ choices of ellipsis remnants. It is difficult to determine what effect sluicing and BAE might have on the realization of remnants by virtue of being the constructions that they are, and I collapse them for the statistical analysis, due to the low frequency of sluicing (see Table 3). Split questions on the other hand, are characterized by adjacency of the correlate and remnant, and exhibit the highest rate of NP remnants. This pattern could have to do with the fact that for every split question I extracted, the correlate is either an in-situ wh-phrase or an instance of sprouting. The former (with what) is shown in (18) and (19) and the latter (by who) in (29).

(29) A: I’ll have you arrested. B: By who? The sand police? (COCA)

In both cases, the remnants directly follow the correlates and are uttered by the same speakers. Placing two PPs (the correlate and the remnant) side by side may be redundant from the processing point of view: because retrieving a directly adjacent correlate could be easier than retrieving a distant one, a reduction in the semantic and syntactic

---

12 Alternatively, one could assign the informativity score of 0 to all correlates for echo questions. I chose not to do this, because it seems to me that some basic features of the correlate (CAT, animacy, number and the fact that there is a referent) are available even if we have not heard it correctly.
content of the remnant is expected. If a preference for NP remnants turns out to be a unique property of split questions, it will provide support for treating construction type as a significant predictor of the realization of remnants.

Reprise questions also differ from sluicing and BAE in that they favor PP remnants. If one grants that in examples (27) and (28) the referents of the correlates are unclear (the interlocutor has not yet identified the entities being discussed), then these remnants differ from the example of sluicing in (30), from Merchant (2010, ex. 39). Notice that in terms of reference here, the remnant is looking for a member of the set of Balkan languages, but it is established that it is the set of these languages that is under discussion.

(30) A: They want to hire someone who speaks a Balkan language.
    B: Really? Which one?

The distinction I am making here between reprise and non-reprise utterances is meant to separate out unidentifiable referents for correlates from referents that are unresolved (that is, give rise to QUDs), and hence license regular sluicing or BAE. Table 3 gives the distribution of the different construction types with respect to the realization of remnants (PP or NP). Notice that there are only nine instances of sluicing in the data. Given this low frequency and no obvious reason to think that sluicing behaves differently from BAE, I collapsed sluicing with BAE for the statistical analysis.

<table>
<thead>
<tr>
<th>Construction type</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>sluicing</td>
<td>2(22.2%)</td>
<td>7(77.8%)</td>
<td>9(100%)</td>
</tr>
<tr>
<td>BAE</td>
<td>93(30.9%)</td>
<td>208(69.1%)</td>
<td>301(100%)</td>
</tr>
<tr>
<td>split questions</td>
<td>10(17.8%)</td>
<td>46(82.2%)</td>
<td>56(100%)</td>
</tr>
<tr>
<td>reprise questions</td>
<td>30(66.7%)</td>
<td>15(33.3%)</td>
<td>45(100%)</td>
</tr>
</tbody>
</table>

Table 3: Realization of remnants by construction type

From this table, we can see that reprise questions exhibit a distinct preference for PP remnants, unlike the other three constructions. It is possible that this preference follows from the relationship of reprise questions with their antecedents.

It is well established in the literature on clarificational utterances that they may retain arbitrarily long portions of their antecedents, from whole clauses ((2) in example (31)) to phrases ((4) and (6) in (31)) (Ginzburg and Sag 2000, Purver, Ginzburg and Healey 2001, Ginzburg and Cooper 2004, Ginzburg 2012).


For some reprise questions, phonological and syntactic parallelism with the antecedents is required (roughly, echo questions). For others (reference questions), the former may be violated (Ginzburg 2012). The example in (22) violates phonological parallelism,

---

13See Sanford and Garrod (1981) for evidence that distance between an anaphor and its antecedent determines how easy it is to retrieve the antecedent.
as do another forty-four instances of reprise questions in my data. An instance of a reprise question preserving phonological parallelism would be the following counterpart of (22), with Yani in place of who.

\[(32)\] A: Well, there is someone that I have just discovered like in the last couple of years and have you ever heard of Yani?

B: Of Yani?

Notice that PP remnants are maximally parallel to their correlates in terms of syntactic nodes: given that the VPs are missing from the remnants, PPs are the highest projections that contain the correlates. I propose that it is in the nature of reprise questions to unambiguously point to the constituent that has not been parsed or heard correctly, and that a more explicit form, one which preserves larger constituents of the antecedent, better serves this function. The strong preference to retain prepositions in reprise questions is entirely expected, if the clarity of the referent is at stake. Recall from section 3.1 that research on anaphoric reference demonstrates that if the speaker considers an antecedent highly accessible to the hearer, they will shorten the pronunciation of proper names (which are highly informative) anaphoric to that antecedent (Fowler et al. 1997). If, however, the accessibility of an antecedent is deemed low by the speaker, they will lengthen the pronunciation of a pronoun (which is less informative than a proper name) anaphoric to that antecedent (Brennan 1995). These strategies help the hearer resolve the anaphors. Thus, one explanation of the behavior of reprise questions is that speakers select a PP remnant as the "lengthened" counterpart of an NP remnant to provide the hearer with more retrieval cues and hence increase their chances of successfully recovering the remnant’s correlate.

Given the different preferences of these four constructions regarding NP and PP remnants, I entered construction type as one of the fixed effects in the model. I now turn to another of these effects—structural persistence.

### 3.3 Structural persistence

Exposure to certain utterances increases the likelihood with which similar utterances are subsequently generated in discourse. This well-studied phenomenon called priming or structural persistence occurs at different levels of linguistic description: semantics, syntax, morphology, lexis and phonetic form (Meyer and Schvaneveldt 1971, Levelt and Kelter 1982, Kempley and Morton 1982, Tanenhaus et al. 1980, Weiner and Labov 1983, Bock 1986, Hartsuiker and Westenberg 2000, Branigan et al. 2000, and Szmiscsanyi 2005). Structural persistence effects are observed at the level of syntax when speakers repeat syntactic structure as their discourse evolves. These effects persist both within individual speakers (self-priming) and in dialogue. Such a reuse of syntactic structure is argued to facilitate language production by allowing speakers to use structures that they have implicitly learned from exposure to other speakers’ output, as well as their own (MacDonald 2013). For example, Levelt and Kelter (1982: 80) demonstrate experimentally that Dutch speakers reuse the syntax of questions like (33) and (34), which optionally contain prepositions, in their elliptical responses. The response to (33) usually corresponds to the structure of the question by also containing the preposition, as in (35), and the response to (34) corresponds to it by containing no preposition, as in (36).
Aan wie laat Paul zijn viool zien?  
‘Who does Paul allow to see his violin?’

Wie laat Paul zijn viool zien?  
‘Who does Paul allow to see his violin?’

Aan Toos.  
’toss.’

Toos.  
’toss.’

Given such evidence, it has been suggested that syntactic variation is at least in part conditioned on similar structures appearing in prior discourse (Branigan et al. 2000, Szmrecsanyi 2005, Bresnan and Ford 2010). In the case of ellipsis, an antecedent hosting the correlate for the remnant appears in the surrounding discourse and could be echoed in the structure of the remnant.

To control for this effect, I annotated each antecedent for whether it contained a continuous correlate (with the preposition pied-piped along with its object) or a discontinuous correlate (with the preposition either separated from its object or missing altogether). If preposition stranding has applied in an antecedent clause, the correlate consists of a preposition separated from, and following, its object (37). It might seem that for antecedent clauses with preposition stranding, the correlate does not contain the preposition but only its object. As can be seen from Table 4 below, however, PP remnants appear with discontinuous correlates, suggesting that the correlate is the entire PP, not only its object. For convenience, I used the label discontinuous to also code correlates appearing in elliptical clauses as NPs rather than PPs, as in (38). The separation of the preposition from its object or its absence could prime speakers to use NP remnants. In contrast, a continuous correlate hosts a preposition preceding its object and is embedded in one of three environments: a declarative clause (39), an in-situ interrogative (40) or an elliptical interrogative clause (often a sprouting remnant) with preposition pied-piping (41).

A: When you think of a wedding cake, what do you think of?
B: Marriage. (SB)

A: He’s in the army.
B: Which one?
A: Ours. (S)

A: Mostly Sharon got on her bandwagon about Missus Jackson.
B: About who? (SB)

A: We’re products of what?
B: A cultural process. (SB)
I hypothesized that, all else being equal, discontinuous correlates are more likely than continuous correlates to yield NP remnants. Table 4 gives the statistics for the realization of remnants with respect to the realization of correlates.

<table>
<thead>
<tr>
<th>Realization of correlate</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>discontinuous</td>
<td>11(10%)</td>
<td>99(90%)</td>
<td>110(100%)</td>
</tr>
<tr>
<td>continuous</td>
<td>124(41.2%)</td>
<td>177(58.8%)</td>
<td>301(100%)</td>
</tr>
</tbody>
</table>

Table 4: Realization of remnants by realization of correlates

Although neither type of correlate prefers PP remnants, an increase can be seen in the number of PP remnants where the correlates are continuous. Quite unsurprisingly, the fact that English tolerates preposition stranding in nonelliptical clauses is relevant to the realization of remnants in terms of structural persistence effects.

### 3.4 Semantic relationship between V and P

This section explores whether the relationship between ellipsis remnants (and their correlates) and the lexical heads affects the realization of remnants. The processing of prepositional phrases has attracted some attention in the psycholinguistic literature. In particular, the argument status of such phrases has been shown to determine how much time readers spend on them. Clifton et al. (1991) and Boland and Blodgett (2006) show that, for verbs biased toward taking PP arguments, PP arguments are read faster than PP adjuncts in nonelided clauses. Arguments are usually assumed to constitute part of the subcategorization frame of particular lexical heads, while adjuncts can freely occur with a wide range of lexical heads. To illustrate, compare the argument PP in a wallet in (42) with the adjunct PP in a hurry in (43). Both PPs are attached to the verb phrases, but only the verb interest in (42) licenses two arguments, an NP and a PP.

(42) The saleswoman tried to interest the man in a wallet during the storewide sale at Steigers. (Clifton et al. 1991: 255)

(43) The man expressed his interest in a hurry during the storewide sale at Steigers. (Clifton et al. 1991: 255)

Boland and Blodgett (2006) argue that the reason for the differential processing of PPs is that arguments, but not adjuncts, are activated already at the point that the lexical heads that subcategorize for them are processed. This allows speakers to form expectations about the upcoming structure and integrate the expected arguments into the evolving discourse faster than the unexpected adjuncts.

Boland and Blodgett’s proposal is based on the Argument Structure Hypothesis (ASH), which assumes that arguments are lexically specified by their lexical heads but adjuncts are specified by syntactic rules (Boland and Boehm-Jernigan 1998), and hence only loosely dependent on the lexical heads they attach to. Analyzing arguments and
adjuncts in different ways allows us to capture a semantic distinction between them: the meaning of an argument PP is dependent on the subcategorization frame of the head, while the meaning of an adjunct PP is only determined by the meaning of the preposition and remains constant regardless of what lexical head the PP modifies. There is evidence that this distinction between argument and adjunct PPs has consequences for the order of (multiple) postverbal PPs. For instance, Hawkins (2000) demonstrates in his corpus study that PPs whose semantic content is dependent on the verb appear adjacent to it 75% of the time. PPs whose semantic content is independent of the verb appear adjacent to it only 45% of the time. The PP order shown in (44) is more likely than that shown in (45), because the preposition for depends on the verb wait for its meaning (though the verb is independently processable). Hawkins (2000) shows that the effect of semantic dependency on the order of PPs is separate from the preference for computing syntactic nodes over short distances. This difference is visible in the order of PPs whose length (measured in words) is similar but which differ in terms of semantic dependency: for such PPs, the order is determined by the semantic dependency between the lexical head and the two PPs.

(44) The man waited for his son in the early morning.
(45) ?The man waited in the early morning for his son. (Hawkins 2000: 10)

Hawkins' characterization of dependent and independent PPs goes beyond the binary distinction between arguments and adjuncts, and reveals a strong bias toward adjacency of constituents that form semantically coherent units. In much the same way as Boland and Blodgett (2006), Hawkins (2000) assumes that information about semantic and syntactic dependencies between words is specified in the mental lexicon and made available to speakers during sentence processing. As a means to explain the adjacency, Hawkins (1999, 2000, 2004) assumes that certain orders of postverbal PPs can better satisfy the Minimize Domains Principle, given in (46), than others.

(46) Minimize domains (MiD)
The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed (Hawkins 2004: 31).

Hawkins distinguishes here between relations of combination and relations of dependency. The former capture our knowledge of verbal subcategorization and the latter refer to the way sentences are processed, and thus both the need to construct an appropriate mother node (here, a VP) and to process elements that belong together semantically are recognized. According to the MiD, adjacency of a verb and preposition means that they may be processed over a short distance. If the verb and preposition are linked by a dependency relation (they form a processing domain in Hawkins’ terms), simultaneous access to both is required before an interpretation for the entire domain may be computed, and hence their adjacency translates into a processing advantage. If the verb and preposition are independent, they may be processed independently, in much the same way that they would be if they were processed out of context. Thus the processing of such constituents is not affected by whether they are adjacent to each other or otherwise.
The observation that semantic dependency between constituents correlates with their adjacency was first articulated in Behaghel (1932). Since then, support for this correlation has come from phenomena such as heavy NP shift, verb-particle constructions, and preposition placement in filler-gap dependencies (Wasow 1997, Wasow 2002, Lohse et al. 2004, Hoffmann 2011). With respect to verb-particle constructions, Lohse et al. (2004) discuss the ordering of constituents in verb-particle constructions. Applying the same entailment tests as Hawkins does, they find that a semantic dependency between a verb and a particle produces the familiar adjacency pattern such that the particle occupies the immediately postverbal position, as in (47), instead of following the direct object, as in (48). Lohse et al. (2004) also test the independence of the effect of semantic dependency by comparing it with the effect of the length of the direct object NP on the word order, and find that the semantic dependency effect indeed is independent.

(47) They carried out a repair.
(48) They carried a repair out. (Lohse et al. 2004: 245)

As for preposition placement, Hawkins (2004) notes that preposition pied-piping is cross-linguistically robust in filler-gap constructions because it facilitates the identification of the gap. That is, when a PP filler is encountered it provides reliable clues as to its function in the clause as adjunct or object. An NP filler (generated via preposition stranding) is three-ways ambiguous with respect to its function: it could serve as the subject, verbal object, or prepositional object. If the parser misidentifies an NP filler that is a prepositional object as a verbal object, it has to later revise this analysis. Despite the advantage of pied-piping, languages in which verbs and prepositions show signs of lexicalization may favor preposition stranding as a way of minimizing the domain in which a verb and preposition are processed (Hawkins 1999, fn. 15). Along similar lines, Culicover (2013) points out that pied-piping ensures that the constituents of a single phrase (PP) are kept adjacent, while preposition stranding does the opposite. This means that the parser can immediately assign more grammatical features to a complete phrase on encountering it as opposed to having to wait to assign features until it has encountered the complete phrase. But it is a reasonable assumption that, depending on the nature of the semantic relationship between a given verb and preposition, processing pressures might push in favor of keeping the two adjacent through preposition stranding. There indeed is evidence that preposition stranding interacts with whether or not verbs and prepositions are semantically dependent on each other such that the stronger the dependency, the more likely the preposition to strand (Hoffman 2011: 269). The correlation is not perfect, however. Consider some examples in(49)-(52) below, where preposition stranding has occurred without the verbs being dependent on the prepositions or the reverse.

(49) Who does Pat do morning runs with?
(50) How many Canadian elections did you vote in?
(51) Which screen did the stimulus appear on?
(52) Who did Jill grow up around?
This evidence and its analysis in terms of the MID is consistent with the results of Boland and Blodgett’s (2006) in that speakers read faster those constituents that they expect to see upon computing subcategorization frames (or processing domains) for the lexical heads they have already encountered. In brief, the evidence shows that speakers recognize semantic dependencies among constituents. Effects of semantic dependency show up both in shorter reading times for constituents that have semantic links with lexical heads and in adjacency of such constituents and the lexical heads in corpora.

It has been independently shown that the processing of sluicing is also affected by the argument status of the remnant and correlate. Frazier and Clifton (1998) and Dickey and Bunger (2011) found that sluicing remnants are comprehended faster if their correlates are arguments (53) than if they are adjuncts (54), and that this pattern is equally robust in sluicing and sprouting.

(53) The secretary typed something, but she wouldn’t tell me what. (Dickey and Bunger 2011: 69)

(54) The secretary typed somewhere, but she wouldn’t tell me where. (Dickey and Bunger 2011: 69)

This finding replicates the results for nonelided sentences, and may similarly be explained by analyzing argument remnants/correlates as representing entities that are activated more strongly than entities that adjunct remnants/correlates represent, because arguments belong in the subcategorization frame of the licensing verb. This activation boost translates into a processing advantage for argument remnants when they are encountered: argument remnants point to referents that are more salient in the current discourse and, hence, easier to recover.

The evidence discussed thus far suggests that lexical heads subcategorize for entire PPs and not just the prepositions used in them. However, semantically related verbs that assign similar thematic roles to PP arguments differ in terms of what prepositions are selected for as part of these arguments. For instance, note that the related verbs in (55)–(57), from Boland and Boehm-Jernigan (1998: 687), all use different prepositions.

(55) The man was blamed for the crime.

(56) The man was charged with the crime.

(57) The man was accused of the crime.

Boland and Boehm-Jernigan (1998) take these data as support for the position that it is the prepositions rather than the PPs that are subcategorized for by the verbs.

Another strand of evidence suggests that both the phonological form and the semantic content of a preposition are subcategorized for, and hence, must be listed in the lexicon entry for each verb (Wechsler 1995, Pollard and Sag 1994, Boland and Boehm-Jernigan 1998). Consider that the preposition to can be assigned the thematic role of location or recipient, which is determined by the licensing verb. In (58), mail licenses a recipient and pin licenses a location, making the sentence ungrammatical. Note that the prepositional object him is held constant here. Compare (58) with the well-formed example in (59), where both verbs select for the same thematic role.
These data reveal that prepositions are more central to the subcategorization frames of lexical heads than whole PPs are, and that the phonological and semantic constraints that lexical heads place on prepositions are likely to have consequences for the salience of remnants’ correlates. I hypothesized that an argument remnant has a more salient correlate than an adjunct remnant, allowing NP remnants more easily. This is because fewer retrieval cues would be needed to point to a correlate in which the preposition is already salient due to its being activated together with the licensing verb before it is encountered in the antecedent clause.

In addition, constraints on the choice of preposition impact the nature of processing domains found in ellipsis. If a preposition itself can satisfy the selectional restrictions of the head, then the two should constitute a processing domain that excludes the object of the preposition. In contrast, a preposition that is used in an adjunct PP has no semantic link with the lexical head, instead forming a processing domain with its object. On this view, separation of a preposition and its object under sluicing should be dispreferred unless the processing of the preposition depends on simultaneous access to the licensing head located in the antecedent. If this is the case, omitting the preposition from the remnant serves to maintain the semantic affinity between the licensing head and the preposition: the head is absent from the remnant, and so is the preposition. However, the nature of the relationship between a verb and preposition remains somewhat unclear in English.

Formalizing this relationship has proven a notoriously difficult problem. One attempt to do this in terms of a syntactic operation of reanalysis is due to Hornstein and Weinberg (1981). This operation takes a verb and a preposition adjacent to it, turning them into a syntactic unit (complex verb), with the effect that the prepositional object behaves like the object of the complex verb post-reanalysis. This is shown in (60).

(60) [V + [P + POBJ]] → [[V + P] + POBJ]

This operation would then account for adjacency of prepositions and verbs in pseudopassives (61) and in clauses with preposition stranding.

(61) This guitar was sat on by a drunk person at a party.

Reanalysis is argued to be problematic, however, for examples such as (62)–(65) from Baltin and Postal (1996), in which the prepositions behave like elements syntactically independent of the verbs.

(62) Fascism was fought for by Goebbels and then, but, I assure you, only then, against by De Gaulle.

(63) Frank talked to Sandra and Arthur *(to) Sally.

(64) The bridge was flown (both) over and under.
(65) *I argued with about such problems the drivers’ union leader.

In (62), the preposition against is separated from the verb fought, and in (63), the preposition to may not be omitted from the gapping remnant, although the verb is omitted. Similarly, in (64), both coordinated prepositions may be separated from the verb. Finally, the prepositional object in (65) may not undergo heavy NP shift while leaving the preposition with behind. These data question the proposal that a preposition and a verb behave like a single constituent in the syntax.

Syntactic reanalysis is problematic not only for English. Norwegian pseudopassives are often cited as incompatible with reanalysis in that strict adjacency is not required between a verb and preposition, even for what look to be complex verbs (Lødrup 1985, Christensen 1986, Åfarli 1989a,b). This is shown in (66) and (67), from Lødrop (1991), for two kinds of Norwegian pseudopassives.

(66) De må bli passet bedre på.
    they must be looked better after
    ‘They must be looked after better.’

(67) Hvorfor passes de ikke bedre på?
    why are looked they not better after
    ‘Why aren’t they looked after better?’

Here, again we have no evidence that the verbs and the prepositions behave like a complex verb in the syntax.

One might propose in defense of reanalysis that it operates not in the syntax, but in the semantics. Reanalysis as formulated in (60) is intended to capture all verbs and prepositions that participate in (mainly) pseudopassive constructions. For some of these verbs and prepositions, there is a semantic relationship between them such that the meaning of one depends on the meaning of the other, resulting in greater loss of compositionality. For other verbs and prepositions, there is no such relationship. Compare the noncompositional pair of look and after in (68) with sleep and in in (69).

(68) Pat is being looked after well.

(69) All the beds were slept in.

The data cited by Baltin and Postal (1996) do not distinguish between these relationships, but it seems that if some sort of semantic reanalysis is operative at all, it is operative over noncompositional pairs of verbs and prepositions like those in (68) rather than those in (69). If we further accept that reanalysis may operate over multiword items, we can capture verbal idioms and other listed combinations that involve prepositions (go through with, put up with, take exception to, and lose sight of).

That such combinations are semantically cohesive has long been recognized in the literature. One early definition goes back to Kruisinga (1925: II/3, 72), who refers to

\[14\] For the purposes of accounting for the behavior of pseudopassives, a mere distinction between those relationships is clearly insufficient. See Bolinger (1977), Riddle and Sheintuch (1983), and Kuno and Takami (2004) for discussion of nonsyntactic constraints on pseudopassives.

\[15\] See also Huddleston and Pullum (2002: 277) for the fossilization view of such combinations, on which their constituents structure is \([V + [PP]]\), but they have the ability to ‘block the application of syntactic processes which can normally apply to such [V-P] combinations’.
'a combination of two or more words forming a semantic unit which is not identical with the combined meanings of its elements'. Similarly semantic criteria, sometimes combined with syntactic ones, reappear as a means to characterize complex verbs in Poutsma (1926), Quirk et al. (1985), Cruse (1986), Claridge (2000), and Tseng (2000). The point is that their semantics should be part of how we describe these combinations. Brinton and Traugott (2005) state explicitly that combinations of verbs and prepositions have been undergoing the process of lexicalization, the result of which is a gradient of these combinations. The most advanced of these are prepositional verbs, and the least advanced ones are free combinations of verbs and prepositions, where the preposition is not selected by the verb in any way. Brinton and Traugott’s (2005: 96) definition of lexicalization is as in (70):

(70) Lexicalization is the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical.

I will assume that combinations of verbs and prepositions are subject to lexicalization, and that they are located at different points of the lexicalization scale. With this mechanism in place, we can now consider some examples of prepositions and verbs in the context of ellipsis. Let’s take the combination fall for (something), a prepositional verb whose meaning is clearly noncompositional. This type of preposition may be considered semantically empty. In (71), including the preposition in the sluicing remnant, and hence separating it from the verb, which is absent from the remnant, seems quite unnatural. The same is true for BAE (72) and even gapping (73). The idea that prepositions may or may not be omitted from gapping remnants based on their semantic relationship with the licensing head was first articulated by Sag (1976), but its empirical accuracy has not been verified until now.

(71) Ben fell for a scam, but I’m not sure what scam/?for what scam.
(72) A: What did Ben fall for?
    B: A scam/?For a scam.
(73) Ben fell for a scam and Ally the same trick as before/?for the same trick as before.

The infelicity of separating a preposition from a verb forming a fully lexicalized combination extends beyond elliptical constructions. Huddleston and Pullum (2002: 275–6) list a number of constructions requiring that prepositions remain adjacent to verbs; these include filler-gap constructions (relative and interrogative clauses), it-clefts, and coordinated PPs. (74)-(77) illustrate. Thus, while we lack conclusive evidence that all V-P combinations occurring in constructions such as pseudopassives and clauses with preposition stranding behave as single constituents in the syntax, speakers still show a sensitivity to the degree of compositionality that characterizes these combinations.

---

16For a comprehensive classification of English prepositions in terms of their semantic content and form-fixedness, see Tseng (2000).
(74) *Everybody’s talking about the scam for which Ben fell.
(75) *For which scam did Ben fall?
(76) *It was for a scam that Ben fell.
(77) *Ben fell for a scam and for some cheap trick.

These examples reveal that the prepositions cannot be pied-piped with their objects, which makes the degradation associated with PP remnants in (71)-(73) unsurprising. More generally, the examples suggest that English speakers’ perceived bias toward omitting prepositions from ellipsis remnants follows from a large number of lexicalized combinations available in this language. At the other end of the lexicalization scale lie prepositions and verbs that have free status with respect to each other. That is, we can assign distinct meanings to both categories. In (78) and (79), the prepositions with and after have distinct meanings that are independent of the verbs, and both kinds of remnants may occur under sluicing.

(78) Damon goes fishing with a gun, but I’m not sure what gun/with what gun.
(79) They first met during a football game, but I can’t remember which game/during which game.

The middle ground between these extremes is occupied by combinations that have less definitive properties. For example, the preposition may be selected by the verb while still carrying some of its independent meaning. For these combinations, we would predict that PP and NP remnants are available, perhaps with a preference for the latter. (80) and (81) illustrate.

(80) Amy’s been waiting for a message, but I don’t know what message/for what message.
(81) Kim invested in some shares, but I don’t know what shares/in what shares.

It is possible that English owes such semantic dependencies to the emergence of combinations of prepositions and verbs in the Middle English period (c. 1150–1500). This development was the consequence of the process that detached prefixes from Old English (c. 450–1150) verbs and moved them to postverbal position. From that point on it was possible for newly formed combinations to begin to lexicalize into meanings that were semantically nontransparent. English has indeed seen a steady growth of prepositional verbs since the Middle English period (Claridge 2000). It is important to note that the emergence of prepositional verbs is unconnected to the development of preposition stranding in English interrogative and relative clauses and in prepositional passive. The earliest records of preposition stranding predate the rise of prepositional verbs by several centuries, as pointed out by Denison (1985). At the same time that the development of preposition stranding and the development of semantic dependencies are historically separate phenomena, the former can support the latter by increasing the range of environments where adjacency of verbs and prepositions is possible. This in turn is likely to extend semantic dependencies even beyond combinations of verbs and prepositions and to combinations of prepositions and other lexical items, such as nouns.
and adjectives. In the current data set, while 75.6% of the items are combinations of verbs and prepositions, 24.4% host prepositions and lexical items other than verbs. Given these facts, I conclude that the relationship between preposition stranding and the strength of semantic dependencies, which appears to be that of mutual reinforcement, is based on the patterns of language use, but not on a shared origin. It is thus reasonable to interpret semantic dependencies between prepositions and other lexical items (henceforth Xs) as a separate factor to be entered in the model. Admittedly, however, the effect of semantic dependency cannot be straightforwardly attributed to speakers following the MiD in that the relevant domain (VP and its constituents) was already computed at the time that the antecedent was processed. What appears to primarily determine the form that the remnant surfaces in is an avoidance of separating lexical items that belong together semantically.

As a means to measure the strength of these semantic relationships, Hawkins (2000) introduced a two-way entailment test. This is based on the idea that some verbs are processed independently of the prepositions with which they may appear (such a V-P combination entails the verb alone). The examples in (78)-(79), in which the V-P combinations are fully compositional, are exactly the case in point. Other verbs, however, are uninterpretable without prepositions (thus the verbs alone are not entailed), which is true of the noncompositional V-P combination in examples (71) and (72). Just as verbs can be dependent on prepositions, so too can prepositions be dependent on verbs. This type of dependency is seen both in examples (71) and (72) and in (80) and (81). Hawkins (2004) argues that an analysis of semantic dependencies cannot simply be reduced to the standard division into obligatory constituents (complements) and optional constituents (adjuncts). Because many PPs fall into neither the complement nor the adjunct category, they elude standard classification. For example, Hawkins (2004: 113) cites the verb to follow, which has an intuitively stronger semantic link with the PP behind a car (this PP fits well into the lexical entry for to follow in the American Heritage Dictionary: ‘to come, move, or take place after another person or thing in order or time’) than with the PP at great speed, without either of them being an obligatory constituent. I apply these entailment tests as formulated by Hawkins (2000: 242–243) and given in (82) and (83) in coding the current data.

(82) Verb entailment test
If [X V PP] entails [X V], then assign \( V_i \). If not, assign \( V_d \).

(83) Pro-verb entailment test
If [X V PP] entails [X Pro-V PP] or [something Pro-V PP] for any pro-verb sentence listed below, then assign \( P_i \). If not, assign \( P_d \).
Pro-verb sentences: X did something PP; X was PP; something happened PP; something was the case PP; something was done (by X) PP.

In identifying dependent and independent combinations of Xs and Ps, I followed the following procedures. Building on Hawkins (2000, 2004), I introduced three distinct levels of semantic dependency: level 0, level 1, and level 2.\(^ {17} \) These are understood as

\(^ {17} \) This classification scheme is similar to Lohse et al.’s (2004) coding of dependency relations in verb-particle combinations, which also builds on Hawkins (2000). Lohse et al. speculate that classification along these lines may merely be an approximation to still finer levels of dependency possibly involved in such combinations. While this is indeed so, Hawkins’ tests are the finest measure available at the moment.
follows:

- Level 0: semantic independence
- Level 1: one-way semantic dependency, where either $X$ or $P$ depends on the other category.
- Level 2: two-way semantic dependency, where $X$ and $P$ depend on each other

Applying Hawkins’ (2000, 2004) entailment tests, I coded all combinations of $X$ and $P$ for the degree of semantic dependency. I tested the dependence of $X$ on $P$ by removing the entire PP from the antecedent clause, and the dependence of $P$ on $X$ by replacing the actual $X$ with an appropriate proform ($Y$ does/did something, something happens/happened, something is/was done, or something is/was the case). Sample tests appear in (84).

\[(84) \text{ Given the sentence } I \text{ grew up in Dallas, is (1) and/or (2) true?} \]
\[(1) \text{ I grew up.} \]
\[(2) \text{ Something happened in Dallas.} \]

In this example, the PP in Dallas is semantically independent of the verb grew up (both (1) and (2) are true and hence entailed), and hence it would be coded as Level 0. In contrast, example (85) would be coded as Level 1. To see this, note that (85) entails Ben’s waiting, indicating that the verb is independent of the PP for a text message. But the preposition depends on the verb: (85) does not entail that Ben is doing something for a text message.

\[(85) \text{ Ben’s waiting for a text message.} \]

To verify the level of semantic dependency for all items I used dictionary listings (the Merriam-Webster’s Collegiate Dictionary). In addition, I conducted a norming study in which participants recruited through Amazon’s Mechanical Turk, a marketplace interface allowing workers to perform various tasks in return for payment, were asked to provide binary judgments following the pattern shown in (84). This method of data collection is quicker and less expensive than laboratory methods, and its results are comparable to those collected in a laboratory setting (Munro et al. 2010, Schnoebelen and Kuperman 2010, Cable and Harris 2011, Gibson et al. 2011, Sprouse 2011). As a means of minimizing the risk that non-native speakers of English took part in the experiment, I followed two procedures. I set a requirement that all participants have US IP addresses. Before completing the experiment participants were asked to answer two comprehension questions about sentences containing appositives, taken from Harris and Potts (2009). Incorrect answers to one or both of these questions, which non-native speakers find difficult to interpret, led to exclusion of that participant’s data. For each X-P combination extracted from the corpora, I collected ten judgments by native speakers of English who decided whether $X$ and $P$ were dependent on each other for interpretation. 92% of the responses agreed with my original coding. I resolved the cases of disagreement by following dictionary listings.\(^\text{19}\)

\(^{18}\)The proform something is/was the case was used for Xs that were nonverbal.

\(^{19}\)In case that the judgments in Table 4 are not fine-grained enough as measures of the relationship between particular constituents, I included a more nuanced measure in my analysis. Pointwise Mutual Information (PMI), the formula for which is given below, measures the probability with which item B appears together
The data appear in Table 5. Note that the number of PP remnants is much higher for the independent items (Level 0), exactly as expected. Retaining prepositions in these cases does not split semantically cohesive combinations the way it does for the dependent items (Level 1 and Level 2). Both Level 1 and Level 2 items strongly favor NP remnants.

<table>
<thead>
<tr>
<th>Dependency level</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100(46.9%)</td>
<td>113(53.1%)</td>
<td>213(100%)</td>
</tr>
<tr>
<td>1</td>
<td>29(20.9%)</td>
<td>110(79.1%)</td>
<td>139(100%)</td>
</tr>
<tr>
<td>2</td>
<td>6(10.2%)</td>
<td>53(89.8%)</td>
<td>59(100%)</td>
</tr>
</tbody>
</table>

Table 5: Realization of remnants by dependency level of X and P

As a further illustration of this pattern, I include Fig. 1, which shows that the difference in the frequency of PP and NP remnants increases as the dependency level of X and P increases. If we probe the data further, we find that the same pattern is true for prepositions dependent on lexical heads (Fig. 2) and lexical heads dependent on prepositions (Fig. 3). These counts include both Level 1 and Level 2 dependencies. In all three figures, PP remnants are indicated as $P$ and NP remnants as $NoP$.

with item A. This measure was intended to capture further gradience in the data, which follows from different degrees of lexicalization.

\[
PMI(A, B) = \log_2(P(A, B)/(P(A)\times P(B))).
\]  

(1)

The values for PMI increase as the probability of two words occurring together goes up. Thus, PMI reflects different degrees of collocational affinity between constituents, which potentially capture still finer distinctions than the three levels of semantic dependence. I computed PMI for each of the items in the data based on their frequencies extracted from COCA. However, the best fitting model, which I discuss in section 4 did not require that PMI be part of it.
Figure 1: Realization of remnants by dependency level of X and P

Figure 2: Realization of remnants by dependency level of P
A closer look at the data reveals an effect of ellipsis appearing in the antecedent clause. As described in the section on structural persistence, the presence of a preposition in the correlate should in general raise the likelihood of seeing one in the remnant. However, nonelliptical antecedents differ from elliptical ones. Overall, continuous correlates are paired with NP remnants almost half the time (see Table 4). This distribution changes when we take the presence or absence of ellipsis into consideration. Consider first the antecedents in (86)-(87), which host both ellipsis and continuous correlates. Such antecedents seem to prefer PP remnants in BAE (note that these are not reprise questions).

(86) A: That was an election — that was a campaign promise.  
B: *By whom?*  
A: By Barack Obama. (COCA)

(87) A: We were in the room that particular night because we were asked to be in the room.  
B: *By whom?*  
A: By Carla Albright. (COCA)

It is possible that the presence of ellipsis in an antecedent strengthens the effect of structural persistence simply by virtue of producing a PP rather than a full clause. Such a PP is not embedded in any higher projection, and therefore remains activated. This is predicted by the hierarchical model of syntactic priming, where, for example, material embedded in matrix clauses yields stronger effects of priming than material
in subordinate clauses (Branigan et al. 2000, Pickering et al. 2000).

The realization of remnants with respect to the distribution of elliptical and nonelliptical antecedents is shown in Table 6. Notice that while NP remnants are more frequent for elliptical and nonelliptical antecedents alike, the appearance of ellipsis causes a reduction in their frequency.

<table>
<thead>
<tr>
<th>Ellipsis in antecedent</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>ellipsis</td>
<td>69(41.9%)</td>
<td>96(58.1%)</td>
<td>165(100%)</td>
</tr>
<tr>
<td>no ellipsis</td>
<td>66(26.8%)</td>
<td>180(73.2%)</td>
<td>246(100%)</td>
</tr>
</tbody>
</table>

Table 6: Realization of remnants by presence/absence of ellipsis in antecedent

A better insight into the involvement of ellipsis can be gained by breaking the data down into continuous and discontinuous correlates. Table 7 shows the distribution of remnants for continuous correlates, and Table 8 the distribution of remnants for discontinuous correlates.

<table>
<thead>
<tr>
<th>Ellipsis in antecedent</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>ellipsis</td>
<td>64(51.6%)</td>
<td>60(48.4%)</td>
<td>124(100%)</td>
</tr>
<tr>
<td>no ellipsis</td>
<td>60(33.9%)</td>
<td>117(66.1%)</td>
<td>177(100%)</td>
</tr>
</tbody>
</table>

Table 7: Realization of remnants by presence/absence of ellipsis in antecedent (continuous correlates)

<table>
<thead>
<tr>
<th>Ellipsis in antecedent</th>
<th>PP remnant</th>
<th>NP remnant</th>
<th>Total remnants</th>
</tr>
</thead>
<tbody>
<tr>
<td>ellipsis</td>
<td>5(12.2%)</td>
<td>36(87.8%)</td>
<td>41(100%)</td>
</tr>
<tr>
<td>no ellipsis</td>
<td>6(8.7%)</td>
<td>63(91.3%)</td>
<td>69(100%)</td>
</tr>
</tbody>
</table>

Table 8: Realization of remnants by presence/absence of ellipsis in antecedent (discontinuous correlates)

For continuous correlates, where the PP remnants outnumber the NP remnants for elliptical antecedents, the presence of ellipsis in the antecedent strengthens the effect of priming. Discontinuous correlates fail to show any strong sensitivity to ellipsis, which could be due to their less frequent use in elliptical antecedents, at least where sprouting is involved. This is the case in (88)–(89). Notice that only continuous correlates are allowed in the elliptical antecedents.

(88)  A: We were in the room that particular night because we were asked to be in the room.
     B: *Whom?/By whom?
     A: By Carla Albright. (COCA)

It is interesting to note that (86)-(87) somewhat resemble split questions in that an elliptical correlate is immediately followed by a remnant, though this happens across two speakers rather than one. For reasons that I cannot quite explain, split questions have a preference for NP remnants, while BAE of this type does not. Tentatively, I propose that a clear sentence boundary (due to a change of speaker) can make the correlate less accessible and hence, the remnant will tend to be more explicit (a PP instead of an NP).
Based on these data, I hypothesized that the presence of ellipsis in the antecedent plays a role in predicting the realization of remnants.

4 Corpus Model

As will be seen shortly, the statistical model of the sample of 411 items confirms the expected patterns gleaned from the raw frequencies of PP and NP remnants. The model predicts, based on five independent variables, the appearance of NP remnants as opposed to PP remnants. There is an advantage to developing such a model for the current data. The independent variables may intersect so that it is impossible to measure the effects of individual variables. Alternatively, variables may tug in opposite directions, in which case the resultant remnant is presumably the product of the influence of the stronger variable. For example, the correlate in (90) has the informativity score of 0.70 and is discontinuous (due to preposition stranding). If these two effects are additive, the preference for an NP remnant is stronger than it would be if the correlate was less informative, as in (91). But we cannot know this without proper statistical tools.

(90) A: And what division are you in?
B: Fourteen. (S)

(91) A: Reggie Roby, who does he play for?
B: Miami. (S)

On the other hand, a correlate with the informativity score of 0.70 which is also continuous, as in (92), creates tension between choosing an NP remnant and a PP one. Here, the former influence wins out. The logistic regression model (Baayen 2008) developed here offers a fine insight into the influence of the different independent variables when predicting the dependent variable, that is, the appearance of NP remnants.

(92) A: So you went to what college?
B: Uh, Northern State University in Aberdeen, South Dakota. (S)

Below I present the best fitting model of the data. It was identified by means of likelihood ratio tests, which compare models with differing numbers of predictors and select the simplest one that successfully captures the data. This model includes five fixed effects: structural persistence, construction type, correlate informativity, semantic dependencies between lexical categories, and the presence/absence of ellipsis in the antecedent. Fixed effects are diagnostics of the utility of those variables in the model that can be replicated from one data sample to another. But there are also three factors in the dataset that make it unbalanced, such as the corpora from which the data were extracted, the speakers providing the data, and all the different combinations of prepositions and their licensing heads. These were all initially included in the model as random effects, but the first two were found to have an unreliable influence on the data, and so were eliminated from the final model. The outcome of the best fitting model,
which correctly predicts 97% of the data, is reported in Table 9. Note that it includes no interactions between any of the fixed effects, because none played a significant role in predicting the variation in question.

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.41609</td>
<td>0.07739</td>
<td>18.298</td>
<td>0.000000</td>
</tr>
<tr>
<td>DependencyLevel=1</td>
<td>0.19895</td>
<td>0.05447</td>
<td>3.652</td>
<td>0.000318</td>
</tr>
<tr>
<td>DependencyLevel=2</td>
<td>0.25155</td>
<td>0.07159</td>
<td>3.514</td>
<td>0.000538</td>
</tr>
<tr>
<td>Informativity</td>
<td>0.29821</td>
<td>0.12238</td>
<td>2.437</td>
<td>0.015253</td>
</tr>
<tr>
<td>Construction=reprise question</td>
<td>-0.39193</td>
<td>0.06586</td>
<td>-5.951</td>
<td>0.000000</td>
</tr>
<tr>
<td>Construction=split question</td>
<td>0.27213</td>
<td>0.06369</td>
<td>4.272</td>
<td>0.000024</td>
</tr>
<tr>
<td>EllipsisInAntecedent=ellipsis</td>
<td>-0.15017</td>
<td>0.04262</td>
<td>-3.524</td>
<td>0.000475</td>
</tr>
<tr>
<td>Correlate=discontinuous</td>
<td>0.15604</td>
<td>0.04810</td>
<td>3.244</td>
<td>0.001295</td>
</tr>
</tbody>
</table>

Table 9: Output of corpus model

This model makes the following predictions. The intercept estimate of 1.41609 must be adjusted upwards for items whose level of semantic dependency is 1 and 2 as opposed to the base level (0). It must also be adjusted upwards for constructions that are split questions compared to the base level (BAE), for items whose correlates are discontinuous compared to the base level (continuous correlates), and for correlate informativity, which is the only numeric predictor in the model. As can be seen from the p-value column, all of these adjustments are significant at $p < 0.05$. The intercept must be adjusted downwards in two cases: for constructions that are reprise questions compared to the base level (BAE) and for items whose antecedents are elliptical compared to the base level (nonelliptical antecedents). Both of these adjustments are significant at $p < 0.05$. All upward adjustments to the intercept increase the preference for NP remnants, while all downward adjustments decrease the preference for NP remnants.

To better illustrate the model’s predictions about the realization of remnants, I extracted corpus probabilities for the remnants in examples (90)–(92), where the effect of structural persistence was at stake. Although all have a high probability of being realized as NP remnants, the probability of (90) is higher (0.98) than the probability of (92) (0.93), reflecting the influence of the discontinuous and continuous correlates. The probability of example (91) is the lowest (0.86). Of course, informativity and structural persistence are not the only factors involved in these examples, since the three V-P combinations differ in terms of semantic dependence, as well (that is, be is dependent on in, go is dependent on to and vice versa, but play is independent of for and vice versa). Hence, the model advances our understanding of what contribution each factor makes to predicting the realization of remnants.

The outcome of this model provides strong support for the involvement of five factors in predicting the realization of remnants. Notice first that construction type influences the data in the expected way: split questions come with a significantly stronger preference for NP remnants than BAE, while reprise questions come with a significantly stronger preference for PP remnants than BAE. In English, where the overall preference is for NP remnants, the pattern revealed by reprise questions stands out and is shown to be statistically significant. Another factor is the effect of correlate informativity. It is not driven by the correlate’s weight (as measured in words), providing support for the independent status of informativity. This is unsurprising, given
that there are English NPs of equal weight but with different informativity scores, for example, compare *this person* and *this lad*. The former’s informativity score is 0.60, and the latter’s informativity score is 0.80, because it encodes natural gender and age in addition to the six features encoded by *this person* (CAT, number, animacy, humanness, concreteness, and a referent). According to the model, the preference for NP remnants increases with increasing informativity scores. The third factor manifests itself in the way that, compared to nonelliptical antecedents, elliptical ones decrease the preference for NP remnants. This effect was not quantified by an interaction with any of the other factors, suggesting that it is the consequence of the distribution of the data we saw in Tables 6, 7, and 8. Continuous correlates are more commonly found in elliptical antecedents than discontinuous correlates and are shown to exert strong structural persistence effects in these contexts. I now turn to the next two effects, which merit considerable attention because they highlight the unique behavior of English remnants.

I hypothesized in section 3.3 that the availability of preposition stranding in English influences the realization of remnants as a pattern to which speakers are commonly exposed. This hypothesis receives strong support in the model: a discontinuous correlate induces a stronger preference for NP remnants than a continuous correlate. To verify statistically that the inclusion of the type of correlate indeed improves the accuracy of the model of the English data, I developed a model that differed from the best fitting model in Table 9 in not involving the type of correlate factor. Using likelihood ratio tests, I found that the new model was significantly worse than the best fitting model ($p < 0.001$), which confirms that the higher frequency of NP remnants in English is partly attributable to the effects of structural persistence.

The last factor involves combinations of lexical items exhibiting different levels of semantic dependency. If three levels of semantic dependency are distinguished, then the two higher levels lead to a significantly stronger preference for NP remnants than the lowest level. This finding demonstrates that speakers recognize that combinations of lexical items show various degrees of compositionality, and that a sensitivity to degrees of compositionally can be traced not only in the word order of these items, but also in how ellipsis remnants that involve them are realized. As a means to test whether semantic dependencies are indeed critical to accounting for the distribution of ellipsis remnants in English, I removed this factor from the best fitting model, leaving all the other factors in, and developed a new model. The new model was found to be significantly worse than the best fitting model ($p < 0.0001$), providing statistical support for the critical role of semantic dependencies. As the final step in the statistical analysis, I found that if both factors structural persistence and semantic dependencies are removed from the best fitting model, the resultant model is also significantly worse than the best fitting model ($p < 0.0001$).

The reasons why both of these factors render the behavior of English remnants unique are two-fold. First, preposition stranding is typologically rare, since it is only found in English and the Scandinavian languages. This means that structural persistence effects might not arise in a language where correlates are primarily continuous and only rarely discontinuous (if the antecedent is itself elliptical and the correlate is realized as an NP rather than a PP). Second, while all three levels of semantic dependency are present in English, it is only levels 0 and 1 that are commonly found in other languages. That is, many prepositions are processable independently of the lexical
heads, while many may require access to the lexical heads without this relationship being mutual. What is much less common outside of English are fully noncompositional combinations, such as prepositional verbs. The ability to develop prepositional verbs is rather uncommon among the world’s languages. One might assume that this ability is shared by at least some of the Germanic languages. However, the history of none of these languages is as well researched as that of English, and hence, any generalizations about these languages would be speculative by nature, though they may well be correct.

The availability of three levels of semantic dependency has another consequence for the behavior of English. Like the ellipsis alternation, preposition stranding is affected by semantic dependencies between prepositions and their lexical heads (Hoffman 2011). This means that preposition stranding is often available exactly in those contexts where NP remnants are also found, creating an impression of a correlation between preposition stranding and the appearance of NP remnants which is in fact the consequence of both phenomena being affected by the same gradient constraint.

The best fitting model’s excellent accuracy is shown in Fig. 4. Recall that the model correctly predicts 97% of the remnants. The figure graphically represents this fact: the better the curve traces the right-hand corner and bottom border of the ROC space the better the model’s accuracy.

![ROC plot](image)

**Figure 4: Model’s accuracy (ROC curve)**

The current data support the view that the factors affecting the ellipsis alternation reflect gradient rather than categorical constraints on the use of remnants in that all of them get violated in the corpora. For example, even items with the highest level of semantic dependency tolerate PP remnants, as shown by the data in Table 5. Like other instances of syntactic variation (e.g., Bresnan 2007), the patterns found in the distribution of ellipsis remnants thus suggest that speakers’ knowledge of this phenomenon is probabilistic. However, the data considered thus far represent language
production, and hence raise the question of whether other speakers’ judgments about these data would confirm the reality of detailed probabilistic knowledge about which contexts favor which type of remnant. I next turn to speakers’ judgments about the naturalness of a subset of the current data, applying the methodology of Ford and Bresnan (2010).

5 Model of Speaker Ratings

Ford and Bresnan and (2010) tested the intuition that corpus probabilities for constructions appearing in corpus data and speakers’ judgments about the naturalness of the same constructions are correlated. This correlation should show in the way speakers assign points to constructions based on how likely they think these constructions are to be uttered. Ford and Bresnan (2010) used the 100-split task to allow participants a range of numbers that best corresponded to corpus probabilities that their corpus model yielded (these normally range from 1 to 0). Participants were asked to consider two continuations of a story and assign to both a number of points that reflected their assessment of how likely they were to have been uttered. The stories were extracted from the corpus, while their continuations were pairs of the actual continuations attested in the corpus and their made-up counterparts instantiating the relevant alternation.

I followed this methodology in the current study by using data extracted from the corpora as experimental stimuli. For each of twenty-three items I selected, the antecedent was followed by the actual remnant used in the corpora and its appropriate counterpart (arranged in random order). For example, if the actual remnant was a PP, its counterpart was an NP. Participants were asked to assign points to the two alternatives based on their naturalness, given the surrounding context (antecedent).

Sample experimental items appear in (93)-(94).

(93) What age group are they looking at?
A: At eighteen.
B: Eighteen.

(94) Well, so who do you like from that category?
A: From New Age?
B: New Age?

I gathered judgments about these items (ten judgments per item) via Amazon’s Mechanical Turk, following the same procedure as before. Points assigned to the remnants attested in the corpora were then entered in a mixed-effects model comparing speakers’ ratings for these remnants and their corpus probabilities yielded by the best fitting model. I hypothesized that the ratings and corpus probabilities should be quantified by a significant interaction, if speakers are guided by the same factors that influence the corpus data. This model controls for differences among speakers regarding their baseline and the range of points they assign to the alternatives. These were treated as random effects alongside the different combinations of X and P.

\[^{21}\text{All points summed to 100, but any combination was allowed, for example, 21 and 79, or 60 and 40.}\]
\[^{22}\text{Every hit I uploaded on Amazon’s Mechanical Turk contained exactly one experimental item, and no two}\]
effect entered in this model were corpus probabilities for the twenty-three ellipsis remnants extracted from the best fitting model. The output of the model of speaker ratings is a significant interaction between corpus probabilities and participants’ ratings ($p = 0.0005$). Fig. 5 shows the relationship between corpus probabilities and speakers’ ratings for NP remnants. The values for speaker ratings, ranging from 1 to 100, correspond to the corpus probabilities, which range from 0 to 1. This makes it straightforward to evaluate the correlation between the two sets of values: when values for corpus probabilities increase so do values for participants’ ratings. This finding demonstrates that participants’ judgments about the naturalness of PP remnants and NP remnants in context may be reliably predicted from the corpus probabilities for these remnants. Given identical context, speakers recognize the more probable alternatives and are able to express their intuitions about them numerically. Thus the idea that speakers have probabilistic knowledge of the ellipsis alternation is supported by both language production data and judgment data.

Figure 5: Relationship between corpus probabilities and speakers’ ratings for NP remnants

6 Discussion

The evidence discussed in this paper offers insight into the variation between NP and PP remnants in ellipsis. Despite the grammaticality of using either type of remnant in English, both the corpus and experimental data reveal statistically significant patterns indicating that implicit knowledge of the ellipsis alternation involves ”soft” grammatical constraints. The realization of remnants is affected by five constraints, all of which

---
hits were available simultaneously. Hence, it was not necessary to control for the order in which participants assessed the experimental items.
reflect performance preferences whose presence has been supported by independent
research on syntactic variation and language processing.

These five constraints define three performance preferences. The first performance
preference, widely attested in corpus data (Ariel 1990, 2001), captures the human pro-
cessor’s tendency to recover accessible antecedents with less explicit anaphors. We see
this preference at work as an independent factor when speakers select NP remnants
(the less explicit alternative) to recover correlates with higher informativity scores.
Correlate informativity effects are modulated by the number of semantic and syntactic
features the correlate is specified for, but not by its weight as measured by the number
of words. Further, two construction types exhibit distinct behavior that I attribute
to the accessibility of correlates. Due to their clarificational character, reprise ques-
tions usually host low-accessibility correlates, which leads to a high frequency of PP
remnants. Split questions are characterized by the reverse pattern. Here, I propose
that the direct adjacency of the correlates and remnants and the fact that they are
both uttered by a single speaker contribute to the accessibility of the correlates, with
the effect that a less explicit remnant is chosen over a more explicit one.

The second performance preference refers to adjacency of constituents that depend
on each other semantically, and hence form processing domains. Although originally
formulated to explain patterns found in word order (Hawkins 2000, 2004, Lohse et
al. 2004), this preference is also revealed in ellipsis data through avoidance of PP
remnants, if prepositions hosted by their correlates form semantically coherent units
with the lexical heads. Let us consider how semantic dependencies are computed with
PP remnants. There is a clear difference between nonelliptical and elliptical construc-
tions in this respect, because semantic dependencies are computed once in nonelliptical
constructions, but twice in elliptical constructions. They are first computed when the
antecedent is processed and the second time when the remnant is encountered, with the
preposition being the only member of the dependency that is encountered. Therefore
the presence a preposition in the remnant is unlikely to affect efficient recognition of
the relevant semantic dependency in the sense that this dependency is already stored
in the discourse model. The data suggest that what the parser prefers is for all mem-
bers of a semantic dependency to be absent from the remnant, presumably because
the computation of the semantic dependency is partial if only one of its members is
reaccessed at the ellipsis site. The current data show that this preference is strong
enough to override correlate informativity effects. The statistical model did not iden-
tify any interaction between semantic dependencies and correlate informativity, which
would be expected if low-accessibility correlates had a much higher frequency of PP
remnants for level-0 items than for level-1 and level-2 items. This finding makes the
fully testable prediction that languages like English, with a wider range of semantic
dependencies, will exhibit higher frequencies of NP remnants overall.

The final performance preference is a well-known tendency to replicate syntactic
structure seen or heard in previous discourse. Because correlates are PPs whose con-
stituents may be continuous or discontinuous, the remnant can take the form that

---

23 The selection of NP remnants for correlates which are semantically dependent on the lexical heads
and/or vice versa could also, at least in part, be linked to correlate informativity. Such correlates are
activated twice: first, when the lexical heads are encountered, activating the correlates as part of their
subcategorization frame, and second, when the correlates themselves are encountered. As the result, they
become more accessible than correlates that are not activated together with their lexical heads.
corresponds to the form of the correlate. It is in this sense that preposition placement possibilities available in nonelliptical clauses have a bearing on the realization of remnants. The current data provide empirical support for this idea: two environments that increase the rate of NP remnants are antecedent clauses with preposition stranding and elliptical correlates that are themselves NPs. Of these two environments, it is the availability of preposition stranding in nonelliptical clauses that underlies differences among languages with respect to the frequency of NP remnants. For example, while the prepositionless sluice in (95) could function as a correlate in a range of languages, the antecedent in (96) would be available in a subset of Germanic languages.

(95)  
A: I'm going to a wedding.  
B: Whose?  
A: Sally's.

(96)  
A: Whose wedding are you going to?  
B: Sally's.

If we make the assumption that NP remnants are more frequent in languages with preposition stranding, then the cause of their frequency is partly explained by the effects of structural persistence.

As is clear from this discussion, there is ample support for the idea that nonsyntactic constraints are needed, and entirely sufficient, to account for the pattern of usage found with ellipsis remnants. Assuming that NP and PP remnants are always allowed by the grammar of English, we could not explain why PP remnants never appear in some environments or do so only rarely. Nor could we readily explain why there are differences between the use of NP remnants and preposition stranding, if the availability of preposition stranding makes the ellipsis alternation available. Recall from examples (7)–(10) given in the introduction, and repeated here for convenience, that the NP remnants in (97) and (98) are unexpected and so is the degradation in the acceptability of PP remnants in (99) and (100), given that none of these examples have corresponding nonelliptical clauses with acceptable preposition stranding.

(97)  
We will use force under certain circumstances, but we can't say what.

(98)  
This theory is right in some sense, but I'm not sure what.

(99)  
Ben will fall for any trick, let alone an old one/?for an old one.

(100)  
Ben didn't come across old documents but toys/?across toys.

The results of the statistical analysis provide clear answers to these problems. The NP remnants in (97) and (98) have accessible correlates with informativity scores of 0.70, which function to support the omission of prepositions. In examples (99) and (100), the PP remnants are degraded because they conflict with the highest level of semantic...
dependency holding between the verbs and prepositions in addition to the correlates informativity scores of 0.70 (99) and 0.80 (100).

These results have implications for theories of ellipsis and theories of grammar. With respect to ellipsis, they challenge the view that the availability of the ellipsis alternation in a language has a syntactic motivation. If the ellipsis alternation is the byproduct of the preposition placement possibilities available in nonelliptical clauses, as assumed by (Merchant 20001), then the majority of the world’s languages (those without preposition stranding) should not allow it, contrary to numerous counterexamples (see the list in fn. 22). It has been argued that such languages have more than one underlying structure from which to derive sluicing remnants, such as either interrogative clauses or clefts (Szczegielniak 2008, Rodrigues et al. 2009, van Craenenbroeck 2010a,b). In the case of clefts, the remnant’s derivation avoids illicit preposition stranding, as shown in (101) from Spanish.

(101) Juan ha hablado con una chica, pero no sé cuál es la chica con la que ha hablado Juan.  
"Juan has talked with a girl but I don’t know which girl it was that Juan talked with." (Rodrigues et al. 2009: 2)

However, not all of the languages in which the ellipsis alteration has been shown to exist have corresponding cleft structures, thus limiting the generality of the cleft source hypothesis (Vicente 2008, Nykiel 2013a). And even in English, the ellipsis alternation is found in contexts in which no plausible underlying sources can be readily proposed (examples (97)–(100)). The picture of the ellipsis alternation that emerges from the current data is that it exists independently of preposition stranding, but may be affected by its availability. The statistical model helps probe the relationship between the ellipsis alternation and preposition stranding in English. This relationship is shown to be not syntactic, but based on structural persistence effects and the fact that both the ellipsis alternation and preposition stranding are sensitive to semantically dependent categories. This picture is fully consistent with the available typological evidence, as well as approaches which do not assume unpronounced syntactic structure in elliptical constructions (Ginzburg and Sag 2000, Culicover and Jackendoff 2005, Sag and Nykiel 2011).

An alternative explanation for the availability of the ellipsis alternation, which I advocate here, is that it follows from general cognitive mechanisms. None of the three preferences discussed above are specific to elliptical constructions, but have been independently motivated outside of them as means to capture robust distributional patterns observed in corpus and experimental data. Ease of processing and communicative efficiency are recognized as motivation for such preferences, and this includes the Minimize domains principle discussed in section 3.4 (Hawkins 2004, 2014, McDonald 2013). The current data provide strong support for the position articulated in Hawkins (2004, 2014) that usage-based preferences have the potential to explain many properties of grammars which often need to be formulated as syntactic stipulations, while setting such preferences aside as third-factor explanations (see Chomsky 2005) may limit descriptive accuracy of grammars. The usage-based preferences identified in this study help explain why the availability of the ellipsis alternation and the overall
preference for NP remnants over PP remnants are not syntactic facts about English. They can also explain why the existing body of cross-linguistic evidence reveals a distinction between more and less informative correlates with respect to the acceptability of NP remnants. The reported pattern is that NP remnants favor NPs as correlates over indefinite pronouns in Spanish and French (Rodrigues et al. 2009), Polish (Nykiel 2013a, Sag and Nykiel 2011), Czech (Caha 2011), and in Russian (Tanya Philippova, personal communication). The persistence of this pattern across these languages and in English requires no further stipulation beyond the assumption that NP remnants serve to recover accessible correlates. Points of difference among languages lie in the extent to which the effects of semantic dependencies and of structural persistence are identifiable in the distribution of NP and PP remnants.

It is important to stress that some theories of grammar already include usage-based preferences as part of their architecture. One example is Bresnan and Hay’s (2008) notion of gradient grammar, which is consistent with probabilistic and exemplar-based grammars. Probabilistic grammars incorporate gradience seen in natural data by using probability distributions over various syntactic alternatives, with contextual information factored into the calculations (Manning 2003, Hale 2003, Antilla and Fong 2004, Levy 2005, Jäger and Rosenbach 2006). Exemplar-based grammars assume that previously experienced expressions form the basis for forming new ones by analogy, and that no explicit rules are involved in this process (Bod 1998, Bod and Kaplan 2003, Bod 2006). The ellipsis alternation is yet another instance of syntactic variation, suggesting that grammar is quantitative and that the choice of particular syntactic alternatives may be reliably predicted by statistical models of empirical data.

7 Conclusion

I have proposed that five factors can account for the gradience found in the pattern of ellipsis remnants in both corpus and experimental data, an instance of syntactic variation that has not been previously explored. These factors may be subsumed under three performance preferences, which are independently known from research on anaphora, constituent ordering, and language production. The preference that leads to reduction of semantic and morphosyntactic content in anaphors that have accessible antecedents is shown to extend to ellipsis remnants. Also evident in elliptical constructions, as much as in nonelliptical ones, is an avoidance of separating constituents which the parser accesses together as it computes their meanings. Finally, I have offered evidence for structural persistence effects such that the form of the remnant is affected by the form of the correlate. I have suggested that the strength of these preferences in various languages forms a basis for evaluating differences and similarities among them with respect to the availability of the ellipsis alternation. The current data speak in favor of usage-based theories of grammar.

26It is more precise to say that this literature usually frames informativity in terms of the form of the wh-remnant (which-NP or bare wh-phrase), and not the correlate. Given, however, that which-NP phrases are normally paired with informative correlates (Nykiel 2013b), there is nothing riding on this distinction.
References


Hofmeister, P. 2008. The after-effects of linguistic form choice on comprehension. Poster presented at CUNY.


Vicente, L. 2008. Syntactic isomorphism and non-isomorphism under ellipsis. Ms. UCSC.


