Constraints on ellipsis alternation: A view from the history of English

JOANNA NYKIEL
University of Silesia

ABSTRACT

I offer a diachronic perspective on English ellipsis alternation, or the alternation between inclusion and omission of prepositions from remnants under sluicing and bare argument ellipsis. The relative freedom to omit prepositions from remnants has not been stable in English; this freedom is connected to the strength of semantic dependencies between prepositions and verbs. Remnants without prepositions are first attested, but remain less frequent than remnants with prepositions, as late as Early Modern English and gain in frequency following this period. I demonstrate that three constraints—correlate informativity, structural persistence, and construction type—predict ellipsis alternation in Early and Late Modern English. However, predicting ellipsis alternation in present-day English requires semantic dependencies in addition to the three constraints. The constraints can be subsumed under principles of language processing and production (considerations of accessibility, a tendency to reuse structure, and a conventionalized performance preference for efficiently accessing constituents that form processing domains), permitting a unified processing account of ellipsis alternation with cross-linguistic coverage.

This paper explores constraints on the use of prepositions in elliptical constructions as an instance of syntactic variation. It focuses on the possibility of omitting prepositions from two elliptical constructions, shown in the B-sentences in (1) and (2), where the A-sentences serve as antecedents.

(1) A: I'm here for the audition. B: (For) Which audition? sluicing
(2) A: Pat is trying to get away from who? B: (From) Old Mother. BAE

The B-sentences are stranded phrases that are semantically equivalent to full clauses. For instance, (1) could be replaced with the full interrogative clause Which audition are you here for?, and, similarly, (2) could be replaced with the...
full clause *Pat is trying to get away from Old Mother*. If the stranded phrase hosts a
wh-phrase (*which audition* in (1)), the construction is called *sluicing*; if it hosts a
non–wh-phrase (*Old Mother* in (2)), then the construction is called bare
argument ellipsis (BAE). I will refer to all such stranded phrases as *remnants*.

Despite the semantic relationship between the constructions exemplified in (1)
and (2) and corresponding full clauses, prepositions are optional in the former, but
not in the latter. In these constructions, the prepositions *for* and *from* are optionally
omitted from the remnants if they are part of the remnants’ *correlates for the
audition, from who*, which are prepositional phrases (PPs). In the rest of this
paper, I will call a remnant without a preposition an *NP remnant*, and one with a
preposition a *PP remnant*. I will refer to the alternation between the two types of
remnants as ellipsis alternation. This paper asks what mechanisms underlie
variable use of prepositions in the two constructions in question, making ellipsis
alternation available.

The rest of the paper is structured as follows. In the next section, I turn to
previous accounts of ellipsis alternation. The Current Proposal section presents
the hypothesis that the distribution of NP remnants and PP remnants in English
can be explained by diachronic factors. The Data section presents data collected
for this study. The following section discusses the method (regression modeling)
and results. In the General Discussion section, I address the implications of the
diachronic distribution of NP and PP remnants for our understanding of the
constraints on ellipsis alternation both in English and in cross-linguistic terms.
The final section concludes.

P R E V I O U S A C C O U N T S

Constraints on ellipsis alternation have been a much debated issue in theoretical
linguistics. A long-standing deletion-based account of ellipsis connects remnants
to corresponding full clauses. It holds that some constraints on ellipsis alternation
are syntactic, that is, they are categorically determined by the availability of
and (2) from the clauses given in (3) and (4), through movement of the remnant
to clause-initial position and subsequent deletion of the unmoved material (the
deleted material is indicated by strikethrough in (3) and (4)). Note that the
prepositional objects can be separated from the prepositions through preposition
stranding in these sentences, which allows them to move to sentence-initial
position alone.

(3) *Which audition are you here for?*
(4) *Old Mother* Pat is trying to get away from.

Merchant (2001, 2004) argued that languages without preposition stranding in full
clauses place a categorical ban on NP remnants under ellipsis, but languages with
preposition stranding do not. This generalization is known as the preposition-stranding generalization (PSG).

However, the PSG is problematic in English.² Chung, Ladusaw, and McCloskey (1995:273) offered evidence that NP remnants may appear in environments where preposition stranding is unavailable in the corresponding interrogative clauses. Example (5) illustrates.

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

Furthermore, NP remnants appear in environments that lack full sentential sources. Example (6) shows an NP remnant following the phrase let alone, although, as argued in Culicover and Jackendoff (2005), it is not derived from a sentential source. Thus preposition stranding cannot be involved in the generation of these NP remnants, suggesting that the deletion based account is not on the right track.

(6) Pat can’t take care of himself, let alone someone else.

A systematic analysis of naturally occurring NP and PP remnants from present-day English (PDE) reveals two main patterns. First, NP remnants are preferred over PP remnants (67.2% of all remnants are NP remnants), which appears to be a cross-linguistically rare pattern. Second, four nonsyntactic constraints on the use of prepositions successfully capture the full range of the data shown (Nykiel, 2014a).³

First, construction type influences the choice of remnant, where the label construction type captures different functions of the remnants. Within the two main construction types of sluicing and BAE, two more types can be distinguished: reprise questions and split questions. Reprise questions are clarificational utterances: they ask for clarification of part of the previous clause, either because it has been misheard or because its referent remains unidentified by the hearer,⁴ as in examples (7) and (8) from the Switchboard corpus (Nykiel, 2014a:10).

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

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

It is easy to see that reprise questions cut across sluicing and BAE: example (7) instantiates sluicing and (8) instantiates BAE. In the rest of this paper, the terms sluicing and BAE only refer to utterances that are not reprise questions. In terms of ellipsis alternation, sluicing and BAE are more often realized as NP remnants than as PP remnants, which is the reverse of the pattern observed in reprise questions.
Split questions are interrogative clauses followed by tentative answers, all uttered by a single speaker, as in example (9) (from Nykiel, 2014a:11). The initial interrogative clauses in split questions are often elliptical or in-situ interrogatives. Note that in (9), the two utterances by speaker B constitute the split question construction of interest to us.

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

Split questions are much like BAE, except that only one speaker is involved. Of all four construction types, split questions exhibit the strongest preference for NP remnants. I will return to this preference when discussing correlate informativity.

The second constraint on ellipsis alternation is linked to the form of the correlate. Speakers’ choice of remnant follows from structural persistence, which refers to a tendency to reuse a certain linguistic form that has appeared in the surrounding discourse. Structural persistence has been shown to be operative at different levels, from semantics to phonetic form (Bock, 1986; Bresnan & Ford, 2010; Ford & Bresnan, 2013; Hartsuiker & Westenberg, 2000; Kempley & Morton, 1982; Levelt & Kelter, 1982; Meyer & Schvaneveldt, 1971; Szmrecsanyi, 2005; Tanenhaus, Flanigan & Seidenberg, 1980; Weiner & Labov, 1983). Particularly relevant to the present argument is Levelt and Kelter’s (1982:80) evidence that Dutch speakers reuse the structure of interrogative clauses when responding to them, as in (10). In (10a), speaker A’s question includes an optional preposition, which tends to be repeated in the response. In (10b), speaker A’s question lacks the preposition, and there is a tendency for the response to lack it, as well.

   to whom lets Paul his violin see to Toos
   “A: Who does Paul allow to see his violin? B: Toos.”

   whom lets Paul his violin see Toos
   “A: Who does Paul allow to see his violin? B: Toos.”

The pattern observed in Nykiel’s (2014a) data is that a contiguous PP correlate tends to appear with a PP remnant, but a PP correlate where the two constituents are noncontiguous (as seen, for example in interrogative clauses with preposition stranding) favors an NP remnant. Example (11a) illustrates a noncontiguous PP correlate, while (11b) illustrates a contiguous one (the correlate is the elliptical PP About what?). One way to interpret these data is that in both examples, speaker B reuses the form of the correlate they just heard when uttering the remnant: the noncontiguous PP correlate promotes the NP remnant in (11a), and the contiguous one promotes the PP remnant in (11b).

The third constraint involves the informativity of a remnant’s correlate, that is, its semantic and syntactic content (Hofmeister, 2007, 2008, Hofmeister, Jaeger, Sag, Arnon, & Snider, 2007; Hofmeister and Sag, 2010; Hofmeister, Jaeger, Arnon, Sag, & Snider, 2013). Consider the instances of BAE in (12) and (13). The correlate containing the interrogative pronoun who in (12) is less informative than that containing the NP what state in (13), because the latter picks out a set of U.S. states (as opposed to, for example, a less specific set of inanimate entities), while the former picks out a set of human entities.

(12) A: Who is Pat eating out with tonight? B: Christie.
(13) A: What state are they in? B: California.

Similarly, in sluicing, a more informative correlate contains an NP and a less informative one, an indefinite pronoun. For instance, the PP correlate on our list in (14) contains an NP and hence carries more semantic and syntactic information about its referent than the correlate from someone in (15), which hosts an indefinite pronoun.

(14) A: I’m putting you on our list. B: What list?
(15) It’s clearly stolen from someone, but I don’t know who.

More specifically, informativity captures differences among phrases in terms of the number of semantic features they are specified for (Nykiel, 2014a). It builds on the notion of inherent accessibility, which is determined by such semantic features as animacy, number, concreteness, and referentiality, where animate, singular, physical, and referential entities are inherently more accessible than their opposites (Jaeger & Wasow, 2008; Prat-Sala & Branigan, 2000). Prat-Sala and Branigan (2000) showed that inherently more accessible entities influence word order choice by typically appearing in prominent sentence positions, such as the subject position. However, what matters for inherent accessibility is not just the manner in which phrases are specified for certain features, but also the number of features they are specified for. Research on memory retrieval demonstrates that phrases specified for more features receive stronger mental representations than phrases specified for fewer features. This boosts their availability for future retrieval from memory (Ariel, 1990, 2001; Craik & Lockhart, 1972; Gallo, Meadow, Johnson & Foster, 2008; Hofmeister, 2007, 2008, 2011; Hofmeister et al., 2007; Hofmeister et al., 2013).

Nykiel (2014a) found that ellipsis alternation shows higher frequencies of NP remnants for correlates containing phrases specified for more features (more informative correlates) than for correlates containing phrases specified for fewer features. This distribution fits in well with the idea that more informative correlates are more accessible. Such correlates are easier to retrieve, and hence, remnants typically lack prepositions, which can be assumed to provide unambiguous cues to the correlate’s syntactic category (PP) and grammatical function.6 It is worth noting here that the high frequency of NP remnants in split
questions, one of the construction types discussed previously, can be explained by high accessibility of the correlate, arguably due to a single speaker uttering both the correlate and the remnant and thus knowing exactly what the remnant serves to retrieve. PP remnants, on the other hand, are expected in a context in which reliable retrieval cues are deemed necessary for reaccessing the correlate. Reprise questions seem to be one such context, because the remnant must pick out, as efficiently as possible, the correlate whose accessibility is reduced due to being misheard or not processed correctly.

The preference for using NP remnants with informative correlates is robust in languages without preposition stranding, but is usually given a syntactic explanation that avoids violating the PSG (Rodrigues, Nevins & Vicente, 2009; Szczegielniak, 2008; van Craenenbroeck, 2010; Vicente, 2008). The argument is that an underlying cleft-like construction serves as the source for an NP remnant in sluicing just in case an interrogative clause is unavailable in which the preposition can strand and delete with the rest of the clause, leaving behind the remnant. In (16) (Rodrigues et al., 2009:2), the correlate includes an informative phrase, the NP una chica, and the cleft-like source does not require preposition stranding, which Spanish would not allow.

(16) 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 not know which is the girl with the que ha hablado Juan.
that has talked Juan’

‘Juan has talked with a girl but I don’t know which girl it was that Juan talked with.”

One problem with such a multiplication of sources for sluicing remnants is that it misses the generalization that correlate informativity affects the realization of remnants across different languages. A further problem is that underlying cleft-like constructions are not the sources for sluicing in languages such as English (Merchant, 2001) and Polish (Nykiel, 2013), and yet, the correlation of correlate informativity and the appearance of NP remnants is evident across elliptical constructions. An alternative analysis is that the effect of informativity is rooted in memory retrieval, as discussed, and hence generalizes to all languages irrespective of their syntax. I pursue this alternative analysis further in this paper.

The final constraint on ellipsis alternation is that combinations of prepositions and other lexical categories (verbs, nouns, or adjectives) showing semantic dependencies favor NP remnants. Such dependencies reflect the extent to which pairs of lexical categories can be processed independently of each other and were originally introduced in place of the distinction between verbal complements and adjuncts by Hawkins (2000, 2004) in his studies of word order. Hawkins (2004) argued that the notion of semantic dependencies accurately captures links between intransitive verbs and PPs that are listed in their lexical entries, although they are not the verbs’ complements. For instance, the verb consist requires the PP of something, but the verb wait optionally combines with the lexically listed
PP for something. Such a lexically listed PP is more adjunct-like than complement-like in character but depends on the verb for its interpretation. This is illustrated in example (17): the first PP for his son is dependent on the verb, but the second PP in the early morning is independent of the verb; the verb is independent of both PPs.

(17) The man waited for his son in the early morning. (Hawkins, 2004:114)

To see these dependencies, consider that (17) can still be parsed if we omit both PPs, indicating the verb’s independence (The man waited). While the second PP can be assigned a temporal interpretation independently of the verb wait (The man sang/yelled at me/biked in the early morning), the first PP has no coherent interpretation without it (The man sang/yelled at me/biked for his son8). Hawkins (2000, 2004) found that the preferred order in English is for the dependent PP to precede the independent one. Further examples of semantically dependent and independent combinations will be provided. In (18), the verb look is independent of the PP at what age group, while the PP depends on the verb for its interpretation. In (19), the verb come up is uninterpretable without the PP with the hottest hors d’oeuvres and so is the PP without the verb, and in (20), neither the verb brought up nor the PP on the Pine Ridge Reservation in South Dakota are semantically dependent on each other. Nykiel (2014a) found that an NP remnant is more likely to occur in a context where semantic dependency holds, as in (18) and (19), than in (20).

(18) What age group are they looking at?
(19) My sons-in-law are good cooks and enjoy cooking and it kind of gets to be a contest to see who can come up with the hottest hors d’oeuvres.
(20) I was brought up on the Pine Ridge Reservation in South Dakota.

The sensitivity of ellipsis remnants to semantic dependencies of this sort points to the involvement of principles of language processing. Hawkins (2004) argued that prepositions form processing domains with lexical categories that are semantically dependent on them (or vice versa), and must be accessed together with them; separation of such domains in nonelliptical clauses comes at a processing cost. Nykiel (2014a) offered evidence that speakers also avoid separating such domains in elliptical clauses. NP remnants preserve processing domains for pairs of semantically dependent prepositions and verbs (neither the preposition nor the verb is present in the remnant), while PP remnants separate them (the preposition but not the verb is present in the remnant). For pairs of semantically independent prepositions and verbs, where prepositions are more strongly associated with their objects, only PP remnants preserve processing domains.

These four nonsyntactic constraints predict that ellipsis alternation is available irrespective of the syntax of a given language (contrary to the deletion-based account of ellipsis). However, what remains unclear is why English has a
pronounced preference for NP remnants over PP remnants, while other languages
in which ellipsis alternation has been investigated do not.

CURRENT PROPOSAL

In this section, I propose that one of the four nonsyntactic constraints on ellipsis
alternation, namely semantic dependencies, has strengthened over time, leading
to the preference for NP remnants over PP remnants, which we see in PDE.
These nonsyntactic constraints and the diachronic change affecting them can
more accurately account for the distribution of NP and PP remnants in English
(and in other languages) than syntactic constraints can.

Hawkins (2000) pointed out that semantic dependencies are strong in PDE,
which Nykiel (2014a) took to be the main reason why NP remnants are more
frequent than PP remnants in English than in other languages.9 I connect the
strength of semantic dependencies between prepositions and verbs to the
emergence of combinations of verbs and prepositions, dating back to the late Old
English period (450–1150) or early Middle English (1150–1500) (Claridge,
2000), and a subsequent increase in their frequency. These combinations range
from free (go to, work with) to partly semantically compositional (invest in, refer
to) to noncompositional (take after, count on), the latter two categories
constituting (semi)prepositional verbs (Huddleston & Pullum, 2002; Quirk,
Greenbaum, Leech, & Svartvik, 1985). By Hawkins’ criterion of semantic
dependence, prepositional verbs are necessarily combinations of semantically
dependent categories due to their partial or complete lack of compositionality.
What we see in English is a shift from weaker to stronger semantic dependencies
as the frequency of prepositional verbs increases. We predict that as the number
of prepositional verbs grows as of Middle English, as argued by Claridge (2000),
it influences ellipsis alternation by boosting the frequency of NP remnants. We
also predict that, if developing strong semantic dependencies is a necessary
(though not sufficient) condition for NP remnants to outnumber PP remnants,
there may be an earlier stage in the history of English when PP remnants are
favored over NP remnants, and that structural persistence, construction type, and
 correlate informativity are sufficient to account for the distribution of the
remnants at that stage. These predictions are not without precedent. Independent
evidence from English shows that usage-based preferences found in syntactic
variation do change over time (Bresnan & Hay 2008; Wolk, Bresnan, Rosenbach
& Szmrecsanyi, 2013).

The emergence of prepositional verbs correlates with the loss of Old English
prefixed verbs (P-V compounds) in late Old English/early Middle English. Prefixed verbs (e.g., besprecan ‘speak about’) were replaced by combinations of
verbs followed by prepositions, many of which have since developed into
prepositional verbs (Brinton & Traugott, 2005; Claridge, 2000; Denison, 1985,
1993; Fischer, 1992; Goh, 2001; Traugott, 1999). One way to account for this
development is to propose structural reanalysis of the sort shown in (21)
(Denison, 1985). The verb comes to form a constituent with the preposition, which was formerly part of the PP.

\[(21) \ [V] [PP] > [V + P] [NP]\]

Although we lack convincing evidence that this is the correct constituent structure for prepositional verbs in PDE (see, e.g., Baltin & Postal, 1996; Hornstein & Weinberg, 1981; Huddleston & Pullum, 2002), it is clear that prepositions have stronger semantic links with these verbs than with their own objects. Focusing on their semantics, Brinton and Traugott (2005) suggested that combinations of verbs and prepositions have undergone the process of lexicalization such that the two elements have become more or less fully unit-like. The result of this process is that in PDE, combinations of verbs and prepositions form a cline, from more lexicalized to less lexicalized.

One diagnostic of whether or not combinations of verbs and prepositions are compositional in PDE is the availability of preposition stranding. According to Huddleston and Pullum (2002), prepositions fall into three categories: unspecified, and two types of specified prepositions, mobile or fixed. Roughly, fixed prepositions strand obligatorily, but mobile and unspecified ones do not. Fixed prepositions and verbs with which they combine exhibit a high degree of noncompositionality, while mobile prepositions have less of a semantic relationship with verbs with which they combine, and unspecified prepositions have none.

From the diachronic perspective, the availability of preposition stranding in English might have promoted a gradual loss of compositionality in combinations of verbs and prepositions. Preposition stranding is known to be sensitive to noncompositional combinations in PDE (Hawkins, 2004; Hoffman, 2011) and could perhaps entrench new ones by widening the range of environments where verbs and prepositions appear adjacent. As Denison (1985) pointed out, however, it is incorrect to assume that the availability of preposition stranding is the cause for the loss of compositionality in prepositional verbs beginning in late Old English/early Middle English. Preposition stranding has persisted in English since the Old English period: it is available in restrictive relative clauses and with R-pronouns and personal pronouns in Old English, but not in interrogative clauses or passives, and spreads to interrogative clauses, wh-relative clauses, and passives as of Middle English (Allen, 1980; Fischer, van Kemenade, Koopman & van der Wurff, 2000; van Kemenade, 1987). If there was a causal link between the availability of preposition stranding and the emergence of prepositional verbs, evidence of their existence should appear much earlier in English. What is likely the case instead is that, with the spread of preposition stranding in Middle English and later on, the pattern of semantic dependencies found in prepositional verbs extends to combinations of prepositions and nouns or adjectives, further increasing the pool of semantically dependent categories (25% of my data to be presented in the next section include prepositions paired with nonverbal categories). Thus it is plausible that a
preposition-stranding language can develop stronger dependencies than a non-preposition-stranding language can, once favorable conditions exist for the development of semantic dependencies.

Although there is little empirical data except for Nykiel (2013), languages other than English are commonly reported to allow NP remnants to a limited extent (see Vicente [2008] for Spanish, Szczegielniak [2008] for Polish, Stjepanovic [2008] for Serbo-Croatian, and Rodrigues et al. [2009] for Spanish, Brazilian Portuguese, and French). NP remnants usually appear with informative correlates, such as that in example (16) (where the NP *una chica* ‘a girl’ is a more informative phrase than an indefinite pronoun *someone* would be in its place). The fact that NP remnants tend to appear with informative correlates in English and across a range of non-Germanic languages suggests that semantic dependencies impact ellipsis alternation in PDE, producing the preference for NP remnants over PP remnants. Given the uniqueness of PDE, it is possible to make specific predictions about ellipsis alternation in the earlier stages of English.

If semantically dependent combinations of verbs and prepositions form gradually over time, then the history of English might provide some insight into the involvement of such combinations in ellipsis alternation. Given that prepositional verbs appear in late Old English/early Middle English, we can isolate a stage of English with weaker semantic dependencies, that is, Old English. By weaker semantic dependencies, I mean a reduced number of combinations of verbs and prepositions (because prefixed verbs still exist) and no combinations that are partially or fully noncompositional. Independent evidence for weak semantic dependencies in Old English comes from the prepositional passive construction (*This sound has been heard of before*). This construction is taken to indicate some degree of semantic dependence between verbs and prepositions that appear in it (Denison, 1985; Fischer et al., 2000; Hornstein & Weinberg, 1981). It is unavailable in Old English, and when it first appears in Middle English, it is found with a small set of combinations of semantically dependent verbs and prepositions, such as *let of* and *tell of* ‘regard’ (Denison, 1985). Old English should thus provide evidence in support, or otherwise, of the hypothesis that ellipsis alternation is sensitive to the strength of semantic dependencies between lexical items as they develop. It is not until the 14th century that the first few instances of prepositional verbs appear and then gain in frequency in the 15th century (Claridge, 2000; Denison, 1985). Claridge’s (2000) data revealed that throughout Early Modern English (1500–1700) prepositional verbs have the frequency of 2400 per million words. Among these verbs, only 17% of all tokens carry maximally noncompositional meanings. Comparing these counts with data for PDE from Biber, Johannsson, Leech, Conrad, and Finegan (1999), we find that the frequency of prepositional verbs is almost 5000 per million words, hence more than twice their frequency in Early Modern English.

I hypothesize that, corresponding to these differences in the frequency of prepositional verbs, there are differences in the frequency of NP remnants. These differences can be understood in terms of the NP remnants to PP remnants ratio...
at different points in time. The cross-linguistic evidence discussed suggests that for PP remnants to have a higher frequency than NP remnants is the more common pattern than the reverse. On the assumption that the availability of prepositional verbs boosts the frequency of NP remnants, the earlier periods of English, and Old English in particular, are likely to differ from PDE by preferring PP remnants over NP remnants. As a means to verify this hypothesis, I explore the ratio of NP remnants to PP remnants over time and gauge how much of these data can be accounted for by structural persistence, construction type, and correlate informativity. I assume that more data can be accurately accounted for by these three factors in Early Modern English than in Late Modern English.

**DATA**

I begin with data from the Old English period, extracted from the Dictionary of Old English Web Corpus. I extracted wh-phrases that were either sluicing remnants or were embedded in wh-interrogative clauses. From among the former phrases, I further extracted those that had PP correlates. From among the latter, I extracted those that were prepositional objects and hence served as PP correlates for BAE remnants. This search yielded remnants that were sluicing remnants, answers to interrogative clauses, or comments on interrogative clauses, including clarificational questions about them (see (8)). While there are instances of both sluicing and BAE in this corpus, I was unable to find relevant NP or PP remnants. Such remnants could have occurred in contexts where appropriate PP correlates are present, but nonelliptical responses are chosen there instead of remnants. Examples (22) and (23) illustrate. Note that the PPs *(Be hwam ‘about whom’ and Ðurh hwæt ‘through what’)* could have served as correlates for remnants.

(22) Seo fostormodor cwæð. Hlæfdige, be hwam cwist þu þæt. Heo hyre antwirde and cwæð. Ær ðam dæge minra brudgifta Ic eom mid manfulre scilde besmiten. “The foster-mother said: ‘Lady, about whom do you say that?’ She answered her and said: ‘Before my wedding day, I am defiled with evil sin.’” (11th c., Apollonius of Tyre, II:15–16)

(23) Ða cwæð se cyng. Ðurh hwæt wast ðu þ. Se man him aswerode & cwæð. Ðeah he hit silf forswige his gegirla hine geswutelað. “Then the king said: ‘Through what do you know that?’ The man answered him and said: ‘Though he would conceal it himself, his garment betrays him.’” (11th c., Apollonius of Tyre, XIV:196–197)

The data harvested from Old English texts thus fail to provide any evidence of ellipsis alternation. An analogous search performed on Middle English data from the Helsinki Corpus and the Middle English Compendium yielded 23 relevant remnants. All of them are realized as PP remnants; some examples follow (the correlates and remnants are in italics).
We be of this other contrey of the partyes of Gaule, quod Agrauadain. (c. 1450–1460, Merlin, XXX:606)

Y haue a word to thee. And Hieu seide, To whom of alle vs? And he seide, To thee (c. 1382–1395, Wycliffe’s Bible, IV.IX:260)

Thou shalt be wedded / vn to oon of tho That han for thee / so muche care and wo But vn to which of hem / I may noght telle (c. 1405, Canterbury Tales, Ellesmere Ms. I:67)

The lack of NP remnants in the Middle English data strongly suggests that they are rare in this period. It would be risky to assume that NP remnants are unavailable in Middle English, given that data from later periods are much richer in dialogue, an environment that is particularly likely to feature ellipsis. However, it is important to note that it is in Middle English that preposition stranding becomes available in nonelliptical wh-interrogative clauses, which serve as sources for sluicing remnants on the deletion-based account of Merchant (2001). The introduction of preposition stranding into wh-interrogative clauses does not correlate with attested examples of NP remnants. The Middle English data thus provide support for the hypothesis that NP remnants remain less frequent than PP remnants as long as semantic dependencies are weak (recall that the first prepositional verbs appear as late as the second half of the Middle English period).

NP remnants are first found in later texts, showing the pattern that is in line with the current hypothesis. The data (including both sluicing and BAE remnants) were extracted from six corpora: the Complete Works of William Shakespeare, the Helsinki Corpus (for the Early Modern English period), A Corpus of English Dialogues (1560–1760), Literature Online (1500–1900), the Proceedings of the Old Bailey (only the 1700s), and ARCHER: A Representative Corpus of Historical English Registers. The search for NP and PP remnants was the same as before. I extracted all wh-phrases and then manually selected those that instantiated sluicing and had PP correlates. I inspected the remaining wh-phrases to select those serving as prepositional objects within wh-interrogative clauses and having BAE remnants as responses to them. The results of this search exclude Yes/No questions of the kind shown in (27), where wh-phrases are not involved.

(27) A: You did not ask for a share in the booty?
B: No.
A: Nor money?
B: No. (1782, Henry Coxill, Theft > grand larceny)

The sample I collected of 351 ellipsis remnants offers insight into both the frequency of NP remnants and the environments in which they have been used over time. I begin by exploring the raw data, but in the Method and Results section, I develop a mixed-effects regression model of these data to verify the results of the preliminary analysis.
Because a portion of these data (54 instances) comes from the corpus of Shakespeare’s plays, it raises the question of whether omission of prepositions is mediated by metrical considerations. Given Shakespeare’s use of iambic pentameter (10 syllables per line) in his rhymed verse, one could assume that a preposition serves to add a syllable when one is needed in a line. Conversely, omission of a preposition would remove an extra syllable from a line. While some such constraint may be operative, it is not categorical in that the majority of clauses containing ellipsis remnants exceed the 10-syllable limit (28) or fall below it (29). There are only nine 10-syllable lines that contain PP remnants and two that contain NP remnants (both with informative correlates). In the case of the NP remnants, it is particularly difficult to determine to what extent meter is involved alongside the informativity of the correlates. For this reason, I chose not to remove these 11 items from the sample. Finally, notice that the number of syllables in the antecedent need not match the number of syllables in the remnant (see (29)), which shows that variation in the use of prepositions is genuinely governed by the factors surveyed in the Previous Accounts section.

(28) SPEED: Nay, I was rhyming: ‘tis you that have the reason.  
VALENTINE: To do what?  
SPEED: To be a spokesman for Madam Silvia.  
VALENTINE: To whom?  
SPEED: To yourself: why, she wooes you by a figure.  
(The Two Gentlemen of Verona, II. I:135–140)

(29) CAMILLO: I am appointed him to murder you.  
POLIXENES: By whom, Camillo?  
CAMILLO: By the king.  
(The Winter’s Tale, I. II:413–417)

To investigate differences in the ratio of NP remnants to PP remnants over time, I divided the corpora into those representing Early Modern English and those representing Late Modern English. The distribution of remnants in Early Modern English (Table 1) shows that, overall, PP remnants (67%) are more frequent than NP remnants. This pattern is robust across four of the five corpora in Table 1. In Late Modern English (Table 2), PP remnants (55%) are also more frequent than NP remnants. Two corpora in Table 2 show that the proportions of PP and NP remnants are reversed with respect to Early Modern English, but the pattern found in Literature Online parallels that found in Early Modern English. Despite the overall preference for PP remnants, the ratio of NP remnants to PP remnants differs significantly between these periods ($\chi^2 = 4.8854, p < .05$).

These patterns suggest that Early Modern English has a more consistent preference for PP remnants over NP remnants than does Late Modern English. Thus the behavior of ellipsis alternation in Early Modern English resembles that found in several other languages (Spanish, French, Polish, Czech, and Serbo-Croatian), where NP remnants are reported to be less frequent than PP remnants
But in Late Modern English, the preference for PP remnants is less apparent. It is relevant to briefly consider these frequency data in terms of usage patterns, to which I turn next.

**Usage**

If we assume that NP remnants derive from full sentential sources with preposition stranding, then we would expect them to be subject to the same prescriptive pressures as preposition stranding. This expectation is not borne out in the data in Tables 1 and 2. At the same time that preposition stranding is gaining in frequency (Early Modern English), NP remnants remain less frequent than PP remnants, and at the time that the use of preposition stranding is suppressed due to the 18th-century normative pressures, NP remnants gain in frequency. There is evidence that the frequency of preposition stranding is high in the late 17th century, but suddenly drops in the late 18th century (Bergh & Seppänen, 2000; Yáñez-Bouza, 2006). This sudden change is argued by Yáñez-Bouza (2006) to be due to different attitudes to language: while the 17th century promotes conversational style, the 18th century comes with restrictions against what is perceived as incorrect usage. And preposition stranding is relegated to incorrect usage. There are also no known restrictions on NP remnants with respect to the level of formality, either in earlier English or in PDE. But preposition stranding is well known to be a characteristic of informal registers.

There is evidence from acceptability ratings that constructions with shared syntactic representations are perceived by speakers in a similar way (Gibson, Piantadosi, Ichinco, & Fedorenko, 2012). It is possible that speakers’ perception of similarities among constructions that share syntactic representations extends beyond acceptability ratings and to contexts of use. One would thus expect the 18th-century restrictions on preposition stranding in nonelliptical clauses to be paralleled by consistent use of PP remnants. However, no such tendency is evident from the data offered in Tables 1 and 2.

The usage patterns found in preposition placement in nonelliptical clauses, however, may have consequences for the strength of the effect of structural persistence. All else being equal, the less preposition stranding there is in the

<table>
<thead>
<tr>
<th>Corpus</th>
<th>PP Remnants</th>
<th>NP Remnants</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shakespeare</td>
<td>30 (56%)</td>
<td>24 (44%)</td>
<td>54</td>
</tr>
<tr>
<td>Helsinki Corpus</td>
<td>4 (67%)</td>
<td>2 (33%)</td>
<td>6</td>
</tr>
<tr>
<td>Archer (1600–1700)</td>
<td>2 (29%)</td>
<td>5 (71%)</td>
<td>7</td>
</tr>
<tr>
<td>Literature Online (1500–1700)</td>
<td>23 (82%)</td>
<td>5 (18%)</td>
<td>28</td>
</tr>
<tr>
<td>Corpus of English Dialogs</td>
<td>77 (71%)</td>
<td>31 (29%)</td>
<td>108</td>
</tr>
<tr>
<td>Total</td>
<td>136 (67%)</td>
<td>67 (33%)</td>
<td>203</td>
</tr>
</tbody>
</table>
antecedents for ellipsis remnants, the less likely speakers might be to omit prepositions from the remnants. The variable preferences for NP remnants in the corpora of Late Modern English could indicate a tension between increasing semantic dependencies and the still high frequency of PP correlates whose constituents are contiguous (in the dataset, there are 297 PP correlates with contiguous constituents compared to 54 PP correlates with noncontiguous constituents). Although the current study merely scratches the surface of these issues, it offers a preliminary usage-based perspective on the relationship of preposition stranding and NP remnants.

**METHOD AND RESULTS**

In this section, I develop a mixed-effects logistic regression model (Baayen, 2008) of the full dataset given in Tables 1 and 2. To do so, I coded the relevant remnants for correlate informativity, structural persistence, and construction type. Given that the current data are historical, it was impossible to maintain consistency across the current and PDE data by coding the remnants for semantic dependency using the tests of Hawkins (2000, 2004). This is not to say that the effect of semantic dependencies on ellipsis alternation cannot be explored directly in historical data, only that it cannot be explored by means of Hawkins’s tests. There is evidence that NP remnants appear with multiword verbs, as identified by Claridge (2000) for Early Modern English, significantly more often than PP remnants, and that this pattern is significantly more robust in Late Modern English than in Early Modern English (Nykiel, 2014b). I chose instead to examine how accurately correlate informativity, structural persistence, and construction type account for the current data, the underlying assumption being that they do so more accurately in Early Modern English than in Late Modern English, reflecting the differing strength of semantic dependencies.

The coding scheme is as follows. As for correlate informativity, I assume that a correlate can encode a maximum of 10 syntactic and semantic features, and the number of features actually encoded is divided by 10, giving the correlate’s informativity score (Nykiel, 2014a). The 10 features and their levels are listed in Table 3.

The syntactic features are standard ones that are or can be expressed on nouns and pronouns in various languages, including those with overt case marking.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>PP Remnants</th>
<th>NP Remnants</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Bailey Trials</td>
<td>2 (10%)</td>
<td>19 (90%)</td>
<td>21</td>
</tr>
<tr>
<td>Archer (1701–1900)</td>
<td>16 (36%)</td>
<td>29 (64%)</td>
<td>45</td>
</tr>
<tr>
<td>Literature Online (1701–1900)</td>
<td>64 (78%)</td>
<td>18 (22%)</td>
<td>82</td>
</tr>
<tr>
<td>Total</td>
<td>82 (55%)</td>
<td>66 (45%)</td>
<td>148</td>
</tr>
</tbody>
</table>
This ensures a wide application of these features, with languages other than English included. The semantic features animacy, humanness, concreteness, and natural gender are basic features familiar from componential analysis, which uses them to capture taxonomic relations among words. Animacy and humanness enter into a one-way implicational relationship with each other, such that humans are always animate, but nonhumans may be either animate or inanimate. This classification takes care of such relationships by keeping the two features separate rather than collapsing them into a single category. I assume that all indefinite and interrogative pronouns encode humanness, some encode both humanness and animacy, and none encode concreteness or natural gender. For instance, the indefinite pronoun *something* and the interrogative pronoun *what* may refer to animate or inanimate entities, all of which are nonhuman, while the pronouns *who* and *somebody* refer to entities that are necessarily both human and animate.

All 10 features differ in terms of which hosts encode them, which is the basis for calculating informativity scores. Some are encoded by every noun and pronoun (e.g., syntactic category, humanness, and referent), but some are encoded only by nouns (e.g., natural gender and concreteness) or by nouns and personal pronouns (e.g., natural gender). For example, the correlates in examples (28) and (29) (*to whom* and *by whom*) both have the informativity score of .5: five features (syntactic category, overt case, animacy, humanness, and referent) are encoded out of the 10 possible ones. In contrast, the correlate in example (30) (*what man... for*) has the informativity score of .7, because seven features are encoded: syntactic category, number, animacy, concreteness, humanness, natural gender, and referent.

(30) HAMLET: What man dost thou dig it for?
FIRST CLOWN: For no man, sir.
HAMLET: What woman, then? (Hamlet, V. I:133–136)

The majority of correlates containing indefinite pronouns in the current data have the informativity score of .4, because four features (syntactic category, animacy, humanness, and referent) are commonly encoded by such pronouns. Correlates containing NPs have the informativity score of .7 or .8 at most, because overt case and grammatical gender are never expressed on English NPs, but most of the other eight features may be. Overall, correlates whose informativity score is .7 or .8 appear with 78% of the NP remnants and with 34% of the PP remnants.

The data were coded for construction type using the four types distinguished in the Previous Accounts section: sluicing, BAE, reprise questions, and split questions. It is relevant to note that split questions, which are the most likely to appear with NP remnants in PDE, in fact appear with 9% of the NP remnants and 8% of the PP remnants, which is unlikely to be a statistically significant pattern. Examples are given in (31) and (32).

and I am sure I really think you deserve a little severity, for not being more afraid."

“Afraid of what, madam?” cried the baronet; “of a young lady’s walking out without me? (1872, Fanny Burney, *Cecilia*)

I coded structural persistence in terms of the form of the correlate. Correlates whose elements were contiguous PPs were labeled as contiguous. Noncontiguous correlates were those expressed as NPs (such as the correlate marked in bold in example (33)) or those in which prepositions were separated from the objects through preposition stranding (34). The data reveal that noncontiguous correlates appear with 31% of the NP remnants but with only 6% of the PP remnants.

(33) A: Have you been about your work?
B: What Work?
A: The Needle work. (1694, French Master, Boyer Abel)

(34) Serj. Jeff: Who did he send it by?
Oliver: The Carrier.
(1680–1719, Trial D4TCOLLE, Stephen Colledge)

The counts for the coding categories are summarized in Table 4.

The picture that emerges from these counts is that correlate informativity is a strong predictor of the appearance of NP remnants. Well over half of all NP remnants are used in the context favoring such remnants (correlates with the informativiy score of .7 or .8). This pattern explains the apparent recalcitrance of the data extracted from the ARCHER corpus for Early Modern English (see Table 1). All of the NP remnants found in these data have correlates with the informativiy score of .7 or .8. Although these remnants outnumber PP remnants, their presence simply follows from the favorable context rather than constituting an exception to the otherwise robust preference for PP remnants. To evaluate the exact role that correlate informativity, construction type, and structural
persistence play in predicting the appearance of remnants, I discuss a logistic regression model of the current data next.

**Results and discussion**

The model predicts the appearance of NP remnants, as opposed to PP remnants, based on three factors (structural persistence, construction type, and correlate informativity) and one interaction term (construction type*correlate informativity). This model was selected by means of likelihood ratio tests and the Akaike information criterion, and then underwent backward elimination of nonsignificant effects. The model includes two random effects: random intercepts for corpus (all the corpora provide samples of varying size) and for pairs of prepositions and verbs/nouns/adjectives (some of which appear several times in the dataset).

The model, in which all three fixed effects make a significant contribution to predicting ellipsis remnants (their $p$ values are <.05), accurately predicts 97.6% of the data (cf. the baseline accuracy of 62.1%). This is represented graphically in Figure 1 (the model’s accuracy is indicated by how well the curve traces the right-hand corner and bottom border of the relative operating characteristic space). There are no multicollinearity issues in the model’s results; the model’s condition number is 8.34.

The outcome of the model is given in Table 5. The model is predicting the appearance of NP remnants. It lists the factors and assesses their contribution to predicting the result. The estimate value specifies upward or downward adjustments that must be made to the intercept. There are four downward adjustments. A PP correlate whose constituents are contiguous requires a downward adjustment to the intercept compared to the base level (a PP correlate whose constituents are noncontiguous), that is, such a correlate reduces the likelihood that an NP remnant will appear. Reprise questions require a

| Table 4. Counts for the coding categories in Early and Late Modern English |
|---|---|---|---|
| Factor | Factor Level | PP Remnants | NP Remnants |
| Informativity score | .4 | 144 (66%) | 30 (22%) |
| | .7 | 63 (29%) | 70 (53%) |
| | .8 | 11 (5%) | 33 (25%) |
| Total | 218 | 133 |
| Construction type | BAE | 161 (74%) | 55 (42%) |
| | Sluicing | 19 (9%) | 35 (26%) |
| | Split questions | 18 (8%) | 12 (9%) |
| | Reprise questions | 20 (9%) | 31 (23%) |
| Total | 218 | 133 |
| Structural persistence | Contiguous | 205 (94%) | 92 (69%) |
| | Noncontiguous | 13 (6%) | 41 (31%) |
| Total | 218 | 133 |
downward adjustment compared to the base level (BAE), and so do sluicing and split questions. The remaining fixed effects all require upward adjustments, which means that, for each factor, the likelihood of seeing NP remnants increases compared to the base level. Statistically significant among all of the adjustments (see the $p$ value column) are those for structural persistence and correlate informativity. In addition, the interaction term also requires upward adjustments to the intercept, one of which is statistically significant: the effect of correlate informativity is stronger for reprise questions than for the base level, BAE.

Table 6 shows the variance of the random effects. Both account for a comparably small amount of variation. Corpus accounts for the distribution of remnants we saw in Tables 1 and 2. Pairs of prepositions and other categories reveal that some pairs attested several times in the data have a bias toward NP remnants (e.g., wait on, talk of).

Several expected patterns are confirmed by the model. First, correlate informativity has the expected effect on ellipsis alternation: the frequency of NP remnants significantly increases with increasing informativity scores. Furthermore, there is statistical support for the effect of structural persistence; PP correlates whose constituents are contiguous lead to a significant drop in the frequency of NP remnants. The involvement of construction type in the current data is somewhat different from PDE in that the behavior of split questions and reprise questions does not differ from the behavior of BAE. However, the interaction between construction type and correlate informativity reveals that

FIGURE 1. Relative operating characteristic curve (model accuracy 97.6%, Somers’s $D_{xy}$ value = .95).
increasing informativity scores cause reprise questions to host NP remnants significantly more often than is the case with BAE. This finding further highlights the relationship between correlate informativity and the realization of remnants. By hypothesis, reprise questions contain elements that are inaccessible, because they have been misheard or parsed incorrectly, and these elements require explicit retrieval cues, such as PP remnants rather than NP remnants. This requirement is significantly weakened by increasing informativity scores.

Additional analyses reveal that the model makes more inaccurate predictions for Late Modern English than for Early Modern English. I extracted 11 items with the highest residuals (amounting to about 3% of the data), where NP remnants were used in place of the expected PP remnants. All of these NP remnants are associated with the context typical of PP remnants: the informativity score of .4 and contiguous PPs as the correlates. Seven of these remnants come from Late Modern English, indicating a difference in the model’s accuracy for the Early Modern English period and the Late Modern English period. When the model is fitted separately to the Early Modern English data and the Late Modern English data, it accurately predicts 98.9% of the former and 96% of the latter. This result fits in well with the expected reduction in the accuracy of the model’s predictions as the strength of semantic dependencies increases in English. We can be confident that the model’s failure to predict these data is partly due to the presence of prepositional verbs: three of the items can be found in Claridge’s (2000) list of prepositional verbs available already in Early Modern English (care for, talk of, and come to (location)).

**GENERAL DISCUSSION**

As predicted, PP remnants are preferred over NP remnants in Early Modern English and much less consistently so in Late Modern English. This result is in line with the hypothesis that semantic dependencies between prepositions and verbs have been developing gradually since Middle English, boosting the frequency of NP remnants. Regression modeling of these data provides statistical evidence that

---

**TABLE 5. Best-fitting model predicting NP over PP remnants in Late and Early Modern English (fixed effects) (n = 351)**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.26</td>
<td>.13</td>
<td>9.74</td>
<td>.000</td>
</tr>
<tr>
<td>StructuralPersistenceContiguousPP</td>
<td>−.32</td>
<td>.06</td>
<td>−5.39</td>
<td>.000</td>
</tr>
<tr>
<td>ConstructionRepriseQuestion</td>
<td>−.36</td>
<td>.22</td>
<td>−1.61</td>
<td>.108</td>
</tr>
<tr>
<td>ConstructionSluicing</td>
<td>−.23</td>
<td>.40</td>
<td>−.57</td>
<td>.571</td>
</tr>
<tr>
<td>ConstructionSplitQuestion</td>
<td>−.19</td>
<td>.33</td>
<td>−.58</td>
<td>.564</td>
</tr>
<tr>
<td>Informativity</td>
<td>.58</td>
<td>.17</td>
<td>3.38</td>
<td>.001</td>
</tr>
<tr>
<td>ConstructionRepriseQuestion:Informativity</td>
<td>.96</td>
<td>.35</td>
<td>2.75</td>
<td>.006</td>
</tr>
<tr>
<td>ConstructionSluicing:Informativity</td>
<td>.64</td>
<td>.58</td>
<td>1.11</td>
<td>.269</td>
</tr>
<tr>
<td>ConstructionSplitQuestion:Informativity</td>
<td>.94</td>
<td>.72</td>
<td>1.31</td>
<td>.193</td>
</tr>
</tbody>
</table>
correlate informativity, structural persistence, and construction type contribute to predicting NP remnants in the two periods under consideration, and they do so with high accuracy.

The effect of correlate informativity opens up the possibility that ellipsis alternation may be couched in terms of memory retrieval. Phrases can become more informative by virtue of being specified for more semantic and syntactic features. This provides them with stronger mental representations that are more accessible for future reference (Craik & Lockhart, 1972; Gallo et al., 2008; Hofmeister, 2011; Suckow, Vasishth, Lewis, & Smith, 2006; Van Dyke & McElree, 2006). An advantage for more informative correlates is simply that at the point where they are retrieved from memory, speakers perceive them as more accessible than less informative correlates and as requiring fewer retrieval cues. The picture that emerges from the diachronic data is that the use of prepositions serves the purpose of manipulating the retrieval cues present in remnants, depending on the accessibility of their correlates. The data strongly suggest that the syntactic category of the remnant (PP vs. NP) is one of the retrieval cues involved in the retrieval of the correlate. This is demonstrated by the great majority of NP remnants actually appearing with correlates with higher informativity scores both in Early and Late Modern English, a pattern whose statistical significance is confirmed by the regression model of the data. Note that correlate informativity explains why NP remnants can appear in examples (35) (repeating (5)) and (36), which are problematic for the syntactic account of ellipsis alternation. Both correlates’ informativity score is .7 (the features encoded are syntactic category, number, animacy, concreteness, humanness, natural gender, and referent), and hence NP remnants are expected in this context.

(35) We will use force under certain circumstances, but we can’t say what.
(36) This theory is right in some sense, but I’m not sure what.

I have also suggested that reprise questions provide contexts in which the correlates are perceived as inaccessible, leading to infrequent use of NP remnants. However, increases in correlate informativity affect the frequency of NP remnants in reprise questions more than in BAE. This pattern shows clearly that increasing informativity scores can override an otherwise robust preference for PP remnants. The relationship between informative correlates and NP remnants is fully consistent with environments in which NP remnants are reported to appear in a host of other languages, as well as the patterns found in

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of Observations</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus</td>
<td>8</td>
<td>.02189</td>
<td>.1479</td>
</tr>
<tr>
<td>Pairs of prepositions and verbs/nouns/adjectives</td>
<td>248</td>
<td>.04288</td>
<td>.2071</td>
</tr>
</tbody>
</table>
PDE data, as discussed in the Previous Accounts section. This makes possible a uniform analysis of ellipsis alternation across languages in terms of memory retrieval, where languages differ in the ratio of NP remnants to PP remnants, but not in the availability of ellipsis alternation.

The accuracy of the model of ellipsis alternation in the current diachronic data helps explain the uniqueness of this alternation in English. Correlate informativity, construction type, and structural persistence accurately account for a large portion of the data in Early and Late Modern English, but, as we saw in the Previous Accounts section, they are insufficient to account for PDE data. This strongly suggests that what distinguishes PDE from the earlier periods are differences in the strength of semantic dependencies, which are in turn linked to the number of prepositional verbs available. These results provide evidence that the strong preference for NP remnants over PP remnants is a fairly recent development in English, and that the constraints on ellipsis alternation are independent of what preposition placement possibilities are permitted by English syntax in nonelliptical clauses (note that preposition stranding has been available in English since the Old English period). By hypothesis, without the ability to develop semantic dependencies between the relevant lexical categories, English would continue to favor PP remnants, much like several languages reported in the literature do.

The current data provide compelling evidence in favor of three principles of language processing and production affecting ellipsis alternation. The first principle captures the effect of correlate informativity on this alternation, which is independently motivated by considerations of accessibility and retrieval of linguistic material, and predicts that NP remnants should be available in most, if not every language. As for the second principle, it is a conventionalized performance preference. Conventionalization can be seen in the frequency of NP remnants, increasing over time, as does the frequency of prepositional verbs. The performance preference that becomes conventionalized is the preference for efficiently accessing constituents that are semantically dependent on each other and hence form processing domains. This preference is robust in constituent ordering in nonelliptical sentences (Hawkins, 2000, 2004; Wasow 2002), as well as in ellipsis alternation in PDE (Nykiel, 2014a). Hawkins (2000, 2004) explained the preference’s presence in nonelliptical sentences by proposing that the parser accesses semantically dependent constituents efficiently if they are adjacent; otherwise, recognition of the relevant processing domain is slowed down. For ellipsis, Nykiel (2014a) suggested that processing domains are accessed incompletely if one constituent of a domain (the preposition) is left in a remnant, while the other is always missing (the verb), and that the parser avoids such configurations. It is clear from the current data that as the strength of semantic dependencies increases in English, so does the frequency of NP remnants. This result fits in well with independent evidence that English has increased its number of interword dependencies (e.g., between a verb and its arguments) over time and that the efficiency of processing such dependencies has been maximized over time by reducing their length (Hawkins, 2012; Tily,
Finally, the effect of structural persistence can be subsumed under the well-documented preference for reusing structure witnessed in the surrounding discourse (Bock, 1986; Levelt & Kelter, 1982; Szmrecsanyi, 2005; Weiner & Labov, 1983). This preference, which is robust in language production behavior, suggests that the availability of preposition stranding in nonelliptical clauses plays a role in ellipsis alternation. If speakers’ choice of remnant is partly influenced by the form of the correlate, a language with preposition stranding produces environments where PP correlates often appear as noncontiguous, promoting NP remnants. There is thus reason to believe that the availability of preposition stranding affects ellipsis alternation, without determining its existence in a language.

CONCLUSION

This paper has offered a diachronic perspective on ellipsis alternation seen in English sluicing and BAE. I have explored the hypothesis that the relative freedom to omit prepositions from remnants has not been stable in English. In particular, I have proposed that this freedom is linked to the strength of semantic dependencies between prepositions and verbs. As data from Old English through Late Modern English show, NP remnants are first attested, but remain less frequent than PP remnants, as late as Early Modern English and gain in frequency later on. I have demonstrated that three constraints (correlate informativity, structural persistence, and construction type) can accurately predict ellipsis alternation in Early and Late Modern English, but predicting ellipsis alternation in PDE requires semantic dependencies in addition to the three constraints. These constraints can be subsumed under three principles of language processing and production, permitting a unified processing account of ellipsis alternation that can capture the distribution of remnants across languages.

NOTES

1. This account stops short of making predictions about when prepositions will be omitted in ellipsis if preposition stranding is available in corresponding nonelliptical clauses.
2. For evidence that the PSG is also problematic from the cross-linguistic perspective, see Nykiel (2013).
3. These data and the diachronic data presented in this paper are fully consistent with direct interpretation approaches to ellipsis, such as Culicover and Jackendoff (2005), Ginzburg and Sag (2000), and Sag and Nykiel (2011), in which ellipsis remnants have no underlying structure and the competence grammar need not place any syntactic constraints on ellipsis alternation. On these approaches, constraints on ellipsis alternation can be understood as following from an interaction of principles of language processing (which I also discuss in this paper) and competence grammar, as suggested in Sag and Nykiel (2011). An anonymous reviewer also pointed out that a catenae-based analysis of elliptical constructions of Osborne, Putnam, and Gross (2012), which includes sluicing and BAE, can successfully account for the data, because it derives ellipsis remnants without movement.
4. For more information on clarificational utterances, see Ginzburg and Sag (2000) and Ginzburg (2012), among others.
5. For more discussion of split questions, see Arregi (2010).
6. For evidence that resolution of sluicing depends on the strength of the correlate’s mental representation and retrieval cues provided by the remnant, see Martin and McElree (2011).
7. There may well be other explanations for this preference, such as the distance between the correlate and remnant. They are immediately adjacent if the antecedent is an in-situ question, but I do not explore here whether in-situ questions are more common in split questions than in BAE.

8. This PP can be assigned its literal meaning for the sake of X, but this is not the meaning it receives when combining with wait.

9. Exactly to what extent the strength of semantic dependencies differs across languages remains to be explored in greater detail. One criterion for assessing the strength of semantic dependencies is whether a language has combinations of verbs and prepositions that are semantically noncompositional (i.e., fully prepositional verbs) or are fossilized in the sense of Huddleston and Pullum (2002), that is, they may never be separated by any syntactic operations.

10. Denison (1985) noted that many new combinations of verbs and prepositions were borrowed or coined in the Middle English period.

11. The appearance of the prepositional passive construction in Middle English has been attributed to a change in case selection by prepositions, such that they govern inherent case (dative or accusative) in Old English, but structural case (objective) in Middle English. However, as Fischer et al. (2000) pointed out, the picture that Denison (1985) painted of prepositional passive’s development in Middle English suggests that the change in case selection is a necessary but insufficient cause.

12. These counts are based on the Lampeter Corpus of Early Modern English Tracts (1640–1740), whose size is 1,172,102 words. I did not use this corpus for the current study due to the text type it contains, which makes ellipsis unlikely to appear.

13. These frequency data are also reported in Nykiel (2014b).

14. There are two further Early Modern English instances of ellipsis remnants, of which one is an NP remnant and one a PP remnant, in the Corpus of Early English Correspondence Sampler.

15. The only exception is the ARCHER corpus, where the majority of the remnants are NP remnants. These NP remnants, however, all have informative phrases as their correlates, and hence, their appearance is unsurprising.

16. Gibson et al. (2012) showed that island violations in two kinds of filler-gap constructions (wh-interrogatives and relative clauses) correlate in terms of acceptability ratings, but there is no correlation between island violations in filler-gap constructions and in constructions without nonlocal dependencies.

17. The issue of how to explore semantic dependencies between verbs and prepositions in historical data is not straightforward. They could, in principle, be recast in terms of the extent to which the presence of a preposition can be inferred from the verb, following, for example, Resnik’s (1996) metric of a verb’s selectional preference strength based on the frequency distribution of its dependents. However, this reasoning quickly becomes circular. If we hypothesize that the same preference underlies word order choice (cf. order of postverbal PPs discussed in Hawkins [2000]) and omission/retention of prepositions in remnants, then we cannot rely on frequency distribution as a diagnostic of semantic dependency.

18. Features of prepositions are not included in this count, because they remain constant while prepositional objects differ in informativity.

19. One additional factor, period, was included in initial modeling to evaluate the contribution of Early Modern English versus Late Modern English, but model comparison performed on the model with the factor period and the model without it showed no reliable difference between models. Although, there was a trend for NP remnants to gain in frequency in the Late Modern English period compared to the Early Modern English period ($t = 1.696$; $p = .141502$).

20. The model was selected and validated using the lme4 and AICmodavg packages.

21. I ignore register as another possible random effect. While most of the corpora are uniform in terms of register, ARCHER, for example, contains texts of different types. What is important for the current study is to establish what the share of NP and PP remnants is in the Early Modern English and Late Modern English periods. I leave it for further study whether or not ellipsis alternation shows register effects.

22. An anonymous reviewer points out that combinations of prepositions and verbs/adjectives/nouns may be introducing some notion of semantic dependencies into the model although their contribution is not tested explicitly. However, this design is necessary, because some combinations appear five times in the sample (e.g., the verb talk and the preposition of), while other combinations appear only once. The model would incorrectly treat the individual items as independent of each other if it did not know that the data were characterized by repeated measures. When this effect is removed from the model, the results remain significant, but the model’s accuracy drops by about 10%.

23. If the condition number is $< 15$, multicollinearity is not a concern.
Another possibility, which I leave open here, is that ellipsis alternation is affected by contextual predictability of remnants. Omission of the relativizer that is known to be more frequent if the context makes the relative clause predictable (Jaeger, 2010). To pursue this possibility, one would have to first distinguish contexts in which remnants are predictable from contexts in which they are unpredictable. One relevant contrast appears to be question-answer pairs (BAE) versus sluicing, as in (i) and (ii).

(i)  A: Which building are you in? B: (In) Classics.

(ii) Pat is in one of these buildings, but I’m not sure (in) which.

REFERENCES


