A corpus-based account of the placement of dependents in phrasal categories: On syntax and performance in English*

Javier Pérez-Guerra
(University of Vigo)

Pérez-Guerra, Javier. 2015. A corpus-based account of the placement of dependents in phrasal categories: On syntax and performance in English. Linguistic Research 32(3), 503-532. This paper investigates the determining factors of the relative ordering of dependents – complements and adjuncts – in Verb Phrases, Noun Phrases and Adjective Phrases in English, in particular the syntactic principle of complements-first (complements next to their heads and adjuncts following the complements) and the performance principle of end-weight. The data show that complements-first is more powerful than end-weight, and that the role of the latter principle is subsidiary, especially when the former is not at work. This study also looks into the connection between compliance with the syntactic principle of complements-first and the type of head which governs the phrasal construction under investigation. The data imply that agreement with complements-first is stronger when the construction contains a verbal or an adjectival head. Finally, this study has a diachronic dimension since it confirms that the syntactic principle of complements-first is especially operative from Early Modern English onwards, that is, after the syntacticisation of word order in English.

(University of Vigo)

Keywords dependent, adjunct, complement, performance, word order

1. Introduction

Head verbs, nouns and adjectives may be followed by other elements which modify or complement the heads within the phrases. In this paper I analyse the distribution

* I am grateful to the following institutions for generous financial support: the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (grant no. FFI2013-44065-P), and the Autonomous Government of Galicia (grant no. GPC2014/060). I am also grateful to the two anonymous reviewers for their comments on an earlier version of the manuscript.
of the constituents which co-occur with, expand or enlarge the governing projections Verb (V), Noun (N) and Adjective (A) in maximal categories of the type Verb Phrase (VP), Noun Phrase (NP) and Adjective Phrase (AP), respectively. The goal is to investigate the determining factors of the relative ordering of such dependents (complements and adjuncts) in VPs, NPs and APs, in particular the principles of complements-first and end-weight.

The minimal pairs under discussion in this study are illustrated in (1), (2) and (3):

(1) a. Now I will deal with the construction in a way which will lead to odd results.
   b. Now I will deal in a way which will lead to odd results with the construction.

(2) a. the author of this book from London [N.A. the author is from London]
   b. the author from London of this book

(3) a. sorry that you missed the train with all my heart
   b. sorry with all my heart that you missed the train

In (1a) the complement with the construction precedes the adjunct in a way which will lead to odd results. The reversed ordering is exemplified in (1b). Similar combinations of complements and adjuncts are shown in the NPs in (2) and in the APs in (3) above.

Two main issues are explored in this study: the rationale of the preference for a specific ordering of complement/adjunct over its reversed counterpart, and the influence of the principles ruling clausal word order on the ordering of dependents in VPs, NPs and APs in English. As far as the forces which account for the distribution of complements and adjuncts are concerned, it is commonly agreed that the order of such constituents is ruled by their relative weight since this eases their processing. An alternative view relies on the fact that complements and adjuncts are

---

1 Even though, on terminological grounds, I am associating processing mostly with the so-called performance principle of end-weight and not with the syntactic principle of complements-first,
ordered in an optimal way, conforming to the constraints of syntactic efficiency. In other words, they are placed in positions where the reader/hearer easily parses and interprets them as such — by default, complements next to their heads and adjuncts following the complements. This second approach assumes that phrases are ordered according to language-internal syntactic principles. In this paper I will contend that the more formal or performance-related forces, such as the preference for long constituents (long phrases or clauses) to be placed in constituent-final positions, have a subsidiary role. This interpretation is in keeping with the view that syntax is responsible for determining a more rigid principled word order, whereas performance should be connected with (more) unprincipled ordering choices made by speakers.

As regards the diachronic dimension of this study, previous corpus-based studies on historical change in word order in English have paid attention to the slots occupied by major clausal constituents (e.g. subjects, as in the summaries in Fischer 1992 and van Kemenade 2012; objects with respect to verbs, as in a number of studies conducted by van der Wurff and colleagues — e.g. van der Wurff 1997; van der Wurff and Foster 1997a, 1997b; Moerenhout and van der Wurff 2005) or to the consequences which information structure has had for the successful design of the clause (Los 2009, 2012; Pérez-Guerra 2012, among others). As already pointed out, in this article I focus on ordering alternatives within phrases and pay attention to two determinants (complements-first and end-weight) which can be investigated on exclusively formal grounds. On the other hand, the ordering of constituents within the VPs, NPs and APs will be related in what follows to the process of syntacticisation which English word order has undergone over the course of its history. In a very sketchy way, the motivation behind the diachronic analysis of the constructions here is the assumption that the consolidation (or syntacticisation) of word order in the English clause by the end of the Middle English period ran parallel in time with the fixation of the ordering of dependents in phrasal categories, which would also follow syntactic criteria. The investigation is based on data retrieved from four historical corpora, which contain material from Old to Late Modern English. It will be argued that the main criterion for the ordering of complements and adjuncts is syntactic in the modern periods and that this finds

processing and syntax should not be taken as opposites since complements-first, that is, the preference for arguments or complements to be ticked off first, is a crucial ingredient of clausal processing.
specific support in the Late Modern English data, when the trends initiated in Early Modern English are consolidated.

The study is organised as follows: in Section 2 I discuss the principles to be evaluated in the empirical analysis. After some preliminary theoretical considerations I describe the database in Section 3. Section 4 deals with the analysis of the data. In Section 5 I interpret the statistical trends found and draw the main conclusions. Finally, I offer a summary of findings and some concluding remarks in Section 6.

2. The theoretical background

The constituents which can precede or follow V, N or A in the VP, NP or AP, respectively, are normally regarded as dependents (see Huddleston and Pullum et al. 2002:24). According to the connection between such dependents and their governing categories, the former are classified as determiners, complements or adjuncts (or modifiers). Since ordering in pre-head position is quite rigid and the possible landing positions for dependents before the heads are severely constrained by the grammar of English, as compared to the post-head periphery, I will here concentrate on the way in which the post-head dependents are organised; in other words, this study will focus on post-head complements and adjuncts.

As regards the ascription of a dependent to the class of complement or adjunct, a number of scholars (see Egan 2008:2, Mair 1990:201, Herbst et al. 2004:xxxiv, among others) have recognised that the definitions of the syntactic relations of complementation and adjunction are undesirably vague in the literature. In this study I will contend that if a dependent is semantically selected or, in generative terms, subcategorised by the head category, it will be classified as a complement; as Matthews (2007:187) notes, a complement is “[a] unit in a construction either required or specifically taken by an individual member of a lexical category”. By

---

2 Huddleston and Pullum et al. (2002:24) use ‘dependent’ as a general function meaning non-head, and distinguish between ‘modifiers’ of nouns (tall girl), of adjectives (very kind), of adverbs (very carefully) and of verbs (They almost suffocated), etc., and ‘adjuncts’, which function as modifiers in clause structure (He behaved annoyingly). Adjuncts, according to Huddleston and Pullum et al. (2002:59), co-occur with predicators (verbal groups) and with the complements of the predicators in a clause. In this paper, following mainstream transformational postulates, I will use ‘dependent’ as the umbrella term for non-head category, which includes complements and adjuncts.
A corpus-based account of the placement of dependents in phrasal categories ... 507

counter, if the semantic connection between the dependent and the head is loose
and, in consequence, the occurrence of the dependent in the phrase is not required
by the referent of the head, then the dependent will be regarded as an adjunct. The
strong semantic link holding between complements and heads is sometimes
accompanied by syntactic dependencies such as lexical restrictions (‘formal
determination’ in Greenbaum et al. 1996:76, or ‘syntactic integration’ in Noonan
2007:101). To give some examples, prepositional complements of V, N or A are
normally introduced by prepositions which are selected by their governing heads;
thus the prepositional complements depending on the verb deal must be introduced
by the preposition with, the complements of the adjective keen are preceded by the
preposition on, and the prepositional complement of the noun compliance requires
the presence of the preposition with. The syntactic marking of the complements may
be based on categorial selection restrictions imposed by the heads; for example, the
verb assume, the adjective certain and the noun hypothesis commonly select
that-clauses as their complements (see, among others, Dixon 2006:15, and Noonan
2007:52). Apart from these co-occurrence restrictions which are explicit in contexts
of complementation, Matthews (1981:124-127) mentions other syntactic factors for
the identification of complements, such as the impossibility of deletion or dropping
(if dropped, the complement is latent4), and the exclusion of complements when the
pattern is saturated, that is, when it contains all the complements which its valency
allows for.

On syntactic, semantic and lexical grounds, in the previous paragraph I have
given support to the distinction between two levels of dependency, as illustrated by
complements and adjuncts. In what follows I will introduce the two main factors
which are said to account for the ordering of complements and adjuncts in a clause.
As already pointed out, in this paper I will investigate the distribution of heads

3 From a cognitive perspective, the impossibility of dropping is explained in, for example,
Bornkessel-Schlesewsky and Schlesewsky (2009:210) as an instantiation of the human parser’s
preference for core to modifying information (their so-called ‘argument-over-adjunct’ preference).

4 That complements can be deleted in surface syntax – and they are latent – implies that deletion
per se is not a good or even valid diagnostic for adjuncthood. To give an example, the sentence
The fact ∆ was rather disappointing illustrates the ellipsis of the complement of the noun fact; that
the complement of fact is latent in this example is corroborated by the impossibility of uttering the
sentence out of the blue without a valid previous context. By contrast, in the sentence The report
was rather disappointing no elements are latent and this implies that, for example, that you wrote
in The report that you wrote was rather disappointing is not a complement but an adjunct.
followed by two dependents, as in examples (1a) [head + complement + adjunct] and (1b) [head + adjunct + complement] below, repeated her for convenience, in order to see whether or not statistical tendencies observed in the data are in keeping with the performance-related and syntactic explanations, already mentioned in Section 1, which I will describe in details hereafter.

(1) a. Now I will deal [with the construction] [in a way which will lead to odd results].
   b. Now I will deal [in a way which will lead to odd results] [with the construction].

According to Hawkins’ (1994, 2004) theory, (1a) is a better performance solution than (1b). For Hawkins, the preference of (1a) over (1b) can be justified in two different ways: by alluding to performance facts, and from a syntactic perspective. First, on performance-related processing grounds, Hawkins (2004) argues that (1a) is better than (1b) because it minimises the amount of structure which has to be processed between the head category and the second dependent. For example, so that the last constituent in the predicate (in a way which will lead to odd results) can be interpreted as a prepositional phrase modifying the head deal in (1a), the human parser has to process the head verb plus the PP with the construction plus the governing category of the prepositional phrase (in), that is, four words (with the construction in). Hawkins’ theory assumes that parsing is complete once the human parser has achieved the correct analysis of the governing categories of all the constituents. Thus, the parsing of (1a) above is successful when the human parser processes the verb (governing category) and its complement (the PP with the construction), and is aware of the syntactic category of the final constituent (prepositional phrase) simply by processing the preposition in. By contrast, in (1b) the human parser has to process ten words (in a way which will lead to odd results with) in order to analyse with the construction as a prepositional phrase complementing deal. Processing ten words is more costly than processing four

---

5 In pairs of examples in which the complement is a noun phrase, the adjunct+complement ordering (Neil donated to the fund ten dollars) is known as Heavy/Focus Noun Phrase Shift in the generative literature (see Rochemont 1978).

6 Rohdenburg (1996) claims that Hawkins’ principles, based on word counts, are simply suggestive rather than exhaustive.
words, so (1a) is better than (1b) in this theory. The performance account thus draws on principles such as Quirk et al.’s (1985:1398) end-weight, which states that long or heavy constituents occupy later positions in the English clause.

Second, the so-called syntactic explanation for the preference of (1a) over (1b) – and also (2a) and (3a) over (2b) and (3b), respectively – is also inspired by Hawkins’ performance theory. He postulates his ‘Minimize Domains’ (or MiD) principle as follows:

Given two or more categories A, B, [...] related by a grammatical rule R of combination and/or dependency, the human processor prefers to minimize the distance between them within the smallest surface structure domain sufficient for the processing of R. (2004:234)

This dependency relation (or grammatical rule) is close to the concept of complementation described above, and MiD states that a construction where a head is followed by a complement and the complement is preceded by additional lexical material is not an optimal solution. Consequently, MiD explains from a syntactic perspective why (1b), (2b) and (3b) illustrate distributions which are not plausible. In later work, Hawkins’ (2007) MiD is replaced with principles such as his ‘Arguments precede X’, where ‘X’ stands for constituents which are not subcategorised by their heads. Such a reformulation of MiD thus approximates to Quirk et al.’s (1985:49-50) ‘complements-first’ principle, according to which complements precede adjuncts.

---

7 Hawkins (1999, 2011), drawing on a number of metrics which calculate the goodness of the different performance solutions affecting the ordering of constituents in the clause, maintains that the smaller the distance between the head and what he calls the ‘constructing word’ of the second dependent (e.g., the preposition in a prepositional phrase), the better/easier the processing of the construction. In Hawkins’ (1999:5) words, “[t]he human parser prefers linear orders that minimize CRDs [Constituent Recognition Domains]” since “[l]ess demands are made on working memory and there is less expenditure of effort in reaching these structural definitions” (2001:7).

8 Lohse et al. (2004) elaborate on the concept of dependency within the framework developed by Hawkins. My complements are comparable to their ‘dependents’ and my adjuncts to their ‘independents’.
3. The data

This investigation focuses on the determining factors that can explain the relative ordering of complements and adjuncts in VPs, NPs and APs from a corpus-based perspective. In addition, the second objective of this study is to determine the existence or otherwise of diachronic variation in English as regards the organisation of dependents in the phrasal categories already mentioned. In this respect, in Sections 4 and 5 I will focus on the distribution of complements and adjuncts in phrasal constituents in the history of the English language. The data have been retrieved from the following corpora:

- for the Old English (henceforth OE) data (c750-1150), the York-Toronto-Helsinki Parsed Corpus of Old English Prose – more than 1.5 million words (see Taylor et al. 2003)
- for the Middle English (ME) period (1150-1500), the second edition of the Penn-Helsinki Parsed Corpus of Middle English (PPCME2) – 1,155,965 words (see Kroch and Taylor 2000)
- for Early Modern English (EModE; 1500-1710), the Penn-Helsinki Parsed Corpus of Early Modern English or PPCEME – 1,737,853 words (see Kroch et al. 2004)
- for (Late) Modern English (LModE; 1700-1914), the Penn Parsed Corpus of Modern British English or PPCMBE – 948,895 words (see Kroch et al. 2010).

All these corpora have a parsed version in which the texts are analysed morphosyntactically using a widely accepted tagset based on Principles and Parameters theory. Since VP complements, for example, are annotated by means of specific tags (e.g. OB, OB1, OB2) in the corpora, it was possible to calculate the frequencies of patterns such as verb+object+adjunct and verb+adjunct+object.9 As regards NPs and APs, since the complements and the adjuncts of nouns and

---

9 The query for the VPs provided examples of verbs followed, in either order, by objects (*OB*) and constituents such as prepositional phrases, adverbial phrases, relative clauses (see Huddleston and Pullum et al. 2002:446, for the analysis of relative clauses as adjuncts), subejuncts, (adverbial) subordinate clauses, etc. The list of examples retrieved by the CorpusSearch query required extensive manual discrimination. The different patterns are illustrated later in this section.
adjectives are not parsed as such in the corpora, I have concentrated on examples containing exclusively either that-clauses or infinitive\textsuperscript{10} clauses as complements, as well as other categories functioning as adjuncts. I have focused on that- and infinitive clauses because they are complements in practically all the NPs and APs in which they occur as dependents. The retrieval process has been undertaken by means of CorpusSearch. Results had to be manually revised since the parsing of the corpora was not entirely correct on a number of occasions.

In what follows, I illustrate some construction types that have been included in my database:

(i) verb-governed contexts:

- verb-complement-adjunct (more specifically, verb-object-adjunct):

(4) I rather fancy he loves nothing in the world so much as one could wish. (CARLYLE-1835,2,297.611) [object + quantifier-phrase adjunct]

(5) Mr. Dickson of Capel-street, brought paintings of birds and flowers basso relievo into great fashion: (OKEEFFE-1826,1,18.201) [object + prepositional adjunct]

\textsuperscript{10} Herbst et al. (2004) also include ing and wh clauses in their taxonomy of complement clauses. To give some examples, in (i) the ing clause terrorists bombing the streets is the complement of the predicator wants, and in (ii) when it’s open for anybody is the complement of the adjective sure:

(i) Nobody wants terrorists bombing the streets.
(ii) I am not sure when it’s open for anybody.

I have not considered examples of ing and wh clauses in my analysis of noun- and adjective-governed contexts because the former are not prototypical examples of complement clauses, since many ing and wh clauses can function as adjuncts, as in (iii) and (iv):

(iii) She said goodbye looking at the sky.
(iv) I modified the example when I realised that it was wrong.
(6) Again, there are those various conditions which cause the symptoms of true 'lumbago,' situated at one time, we suppose, in the muscles of the back, and another time in the thick aponeuroses which cover them. (POORE-1876,173.212) [object + reduced relative clause functioning as adjunct]

(7) But to contrive to see all the boys alone for a minute or two is possible, (BENSON-1908,83.410) [object + adjectival adjunct]

(8) We sell no French wines here, Sir. (GOLDSMITH-1773,60.962) [object + adverbial adjunct]

(9) The post served me just as it did y=r= Losp. (ANHATTON-E3-H,2,211.4, 1690-1695) [object + adverbial clause functioning as adjunct]

(10) Cha. Bring your own weapons, be they what they may – (COLLIER-1835,23.833) [object + subordinate (nonfinite) clause functioning as adjunct]

(11) being accustomed to vary the same sentence different ways, [they acquire, by degrees, a readiness and a copiousness of expression.] (CHAPMAN-1774,211.324) [object + nominal adverbial]

- verb-adjunct-complement. Some of the categories fulfilling the function of adjunct are: quantifier phrases (illustrated in (12)), prepositional phrases (in (13)), APs (in (14)), adverb phrases (in (15)) and nominal adverbials (as in (16)):

(12) neither will I again smite any more every thing living, as I have done. (ERV-OLD-1885,8,20G.321) [quantifier-phrase adjunct + object]

(13) [let them] take after it four ounces of treacle-water, (ALBIN-1736,4.67) [prepositional-phrase adjunct + object]
A corpus-based account of the placement of dependents in phrasal categories ...

(14) he hasn’t cut open the leaves, I see.\(^\text{11}\) (COLMAN-1805,20.58) [adjectival dependent + object]

(15) [and she has invited my mother] to spend there the time of Mrs. F. A.’s confinement, which she seems half inclined to do. (AUSTEN-180X,171.232) [adverbial dependent + object]

(16) and, as the members began to muster, seized one by one all the chiefs of the Presbyterian party. (OMAN-1895,401.492) [nominal adverbial + object]

(ii) noun-governed contexts:

- noun-complement-adjunct:

(17) [There is] a wise saying that nine-tenths of the noble work done in the world is drudgery, which is often misused as if it meant that nine-tenths of the drudgery done in the world is noble work. (BENSON,46.109, 1908) [that-clause complement + nondefining relative-clause adjunct]

(18) [King James sent a Person down to him, with] Offers to mitigate his Fine upon Conditions of ready Payment, to which his Lordship reply’d, that if his Majesty pleas’d to allow him a little longer time, he would rather chuse to play double or quit with him: (CIBBER,44.134, 1740) [nonfinite-clause complement + nondefining relative-clause adjunct]

\(^\text{11}\) The idiosyncratic and, on occasions, idiomatic behaviour of resultative constructions like (14), as acknowledged in Goldberg and Jackendoff (2004), might lead to the inadequacy of explaining these examples by using only syntactic and performance-related factors. In this study I have opted for relying on exclusively the shallow design of the VPs, NPs and APs and for trying to account for the different distributions of dependents in such phrases by means of the syntactic and the performance-driven principles. In consequence, I cannot apply criteria other than complements-first and end-weight. That said, the number of resultative constructions in my data is very small and neither their consideration nor their exclusion alters the overall results.
(19) [he also expressed] an opinion that mulattoes inherited the vices of both races - a maxim which I had heard often enough before, (READE, 225.616, 1863) [that-clause complement + appositive adjunct]

- noun-adjunct-complement:

(20) Had a letter from my dear wife that Sandy was quite well. (BOSWELL-1776, 40.120) [prepositional adjunct + that-clause complement]

(21) Also as hinted, an order to all Smiths to make pikes with their whole soul. (CARLYLE-1837, 1,153.458) [prepositional adjunct + nonfinite-clause complement]

(22) I have thought it best to send through Hammond by wire a suggestion which occurred to me that the conduct of the Spaniards in the case of the Deerhound might afford us a fair plea for holding our hands in the case of the Ironclads - (GLADSTONE-1873, 2,408.646) [relative-clause adjunct + that-clause complement]

(23) I say this with Respect to the impetuous Desire I had from a Youth, to wander into the World; (DEFOE-1719, 202.146) [relative-clause adjunct + nonfinite-clause complement]

(24) [And there was] a feeling by no means uncommon, and very deadly, that India would be lost for ever, and with it all the glory of England. (TROLLOPE, 177.356, 1882) [adverbial + that-clause complement]

(iii) adjective-governed contexts:

- adjective-complement-adjunct:

(25) We are just as certain that our Lord is risen as if we had ourselves witnessed His resurrection; (FROUDE-1830, 2,48.329) [that-clause complement + adverbial]
(26) and we are no more fram’d to talk English or any other Modern Language, than we are to talk Latin: (ANON-1711,5.27) [nonfinite-clause complement + than-phrase adjunct]

(27) Nature, in all her various ways of acting, is not so difficult to be understood, as she is darkened by our hard expressions and obscure way of speaking. (BARCLAY-1743,15.43) [nonfinite-clause complement + as-clause adjunct]

(28) [He told him they were] fully resolv’d to dye for their Country, and ready to fight it out to the last Man, if Occasion requir’d, (HIND,310.144, 1707) [ [[adjective head + nonfinite-clause complement] & [adjective head + nonfinite-clause complement]] + adverbial]

- adjective-adjunct-complement:

(29) Now we see men conscious with a great exultation that they become sons of God. (TALBOT-1901,97.133) [prepositional adjunct + that-clause complement]

(30) more than one of them had Charms sufficient at their leisure Hours, to calm and mollify the Cares of Empire. (CIBBER-1740,55.263) [prepositional adjunct + nonfinite-clause complement]

4. Analysis of the data and preliminary findings

This section deals with the distribution of the dependents (complements and adjuncts) in the VPs, NPs and APs registered in the historical database described in Section 3. The data will show whether the optimal placement of dependents is justified by structural facts (weight or length of the dependents) or by the syntactic principle that complements must (immediately) follow their governing heads (the so-called complements-first principle),\(^{12}\) the latter leading to orderings like

\(^{12}\) Both performance-related and syntactic principles are brought into play when Hawkins (2001:20)
verb/noun/adjective-complement-adjunct. I will also explore the question of whether or not the fixation of clausal word order before the Modern period accords with a change in the distribution of complements and adjuncts in phrasal categories. The methodology is based on periodised, statistically validated frequency counts and the resulting relative proportional distributions.

This section is organised as follows: first, Section 4.1 focuses on the so-called syntactic explanation of the distribution of dependents, not only in the verb-governed contexts (4.1.1) but also in the noun and adjective phrases (4.1.2) retrieved from the corpora. The findings will indicate whether or not the data have conformed to (syntactic) principles such as MiD, ‘complements-first’ or ‘Arguments precede X’ from OE to LModE. Second, in Section 4.2 I discuss the so-called performance-centred explanation, which relies on the maxim of end-weight.

### 4.1 'Complements-first'

#### 4.1.1 Verb-governed contexts and 'complements-first'

This section presents the analysis of the VPs in which the verbs (or verb groups) are followed by complements (objects, in particular\(^{13}\)) and adjuncts in their predicates.\(^{14}\) A terminological note is in order here: the examples in which the head is followed by a complement (object) and the complement (object) is followed by an adjunct will be classified as ‘complement-first’; by contrast, I will use the label ‘complement-last’ when the adjunct precedes the complement (object).\(^{15}\)

---

\(^{13}\) Postulates that “adjuncts will only intervene between a head and a complement [N.A.: syntactic dimension] if they are short [N.A.: performance-driven hypothesis].”

\(^{14}\) Future research might shed light on the behaviour of other distributions in predicates containing complements other than objects, such as prepositional complements (\textit{dispense with-PP, dispense with ing-clause} – see Rudanko 1989, 1999 and 2000 for the diachrony of constructions governed by verbs, adjectives and nouns complemented by \textit{to, ing} and \textit{to ing clauses}).

\(^{15}\) This study is based on data in which the process of linearisation affects only two constituents, namely the complement and the adjunct, in order to ascertain the validity of the basic performance-related and syntactic principles as outlined in Second 2. A consequence of this is the exclusion of two-complement structures such as ditransitive (eg. \textit{give object\(_1\) object\(_2\)}, transitive-prepositional/adverbial (eg. \textit{drive object to-PP/clause, put object adverbial}) and complex-transitive (eg. \textit{drive object object-predicative}) patterns from this account. See, among others, Larson (1988) and Johnson (2006) for generative accounts of the ordering of objects in double-object constructions.

\(^{15}\) For clarification purposes, where the syntactic principle ‘complements-first’ postulates that
A corpus-based account of the placement of dependents in phrasal categories ... 517

Table 1 shows the raw figures and percentages of complement-first/last verb-governed constructions in my database. These findings are also laid out in Figure 1:

<table>
<thead>
<tr>
<th></th>
<th>OE</th>
<th>ME</th>
<th>EModE</th>
<th>LModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>complement-first</td>
<td>8270</td>
<td>17803</td>
<td>7859</td>
<td>13084</td>
</tr>
<tr>
<td></td>
<td>71.54%</td>
<td>72.39%</td>
<td>79.75%</td>
<td>83.53%</td>
</tr>
<tr>
<td>complement-last</td>
<td>3290</td>
<td>6790</td>
<td>1995</td>
<td>2579</td>
</tr>
<tr>
<td></td>
<td>28.46%</td>
<td>27.61%</td>
<td>20.25%</td>
<td>16.47%</td>
</tr>
</tbody>
</table>

Figure 1 shows that most of the examples are complement-first from OE to LModE, and plots the increase in the number of complement-first verbal constructions. Whereas the overall distribution is statistically significant ($p<.0001$, $\chi^2=869.38$), the statistical variation from OE to ME is not significant ($p=0.0949$, $\chi^2=2.79$). By contrast, the increase in the frequency of complement-first constructions in EModE and LModE, with respect to ME, is significant (ME>EModE: $p<.0001$, $\chi^2=200.41$; EModE>LModE: $p<.0001$, $\chi^2=58.49$). The

complements precede adjuncts, a construction is either ‘complement-first’ or ‘complement-last’ depending on the actual distribution of the complements and the adjuncts (complement+adjunct or adjunct+complement, respectively).
findings set out in Table 1, then, indicate that EModE is the pivotal period where the frequencies for complement-first and complement-last orderings in verb-governed constructions undergo variation in the language with respect to previous periods, and that it is in LModE that such statistical tendency is consolidated in the language: whereas more than 75 percent of the examples were complement-last in OE and ME, the ratio increases to 84 percent in LModE.

This finding must be related to the evolution of word order in the history of English. In OE “word order [was] regulated by pragmatic tendencies” (Van Hoorick 1994:53) and the placement of subjects, objects and adverbials was less unproblematic according to the flexible syntactic architecture in this period. The establishment of the distribution of constituents in the clause is assumed to be in progress in ME (see Fischer 1992:371), and it is in EModE that clausal and phrasal English word order is definitively fixed or syntactised. In consequence, subjects, objects and adverbials have a designated unmarked slot in the clause.\(^{16}\) Even though the OE and the ME data are in keeping with the syntactic principle of complements-first, the fact that the drift from a pragmatically-determined to a syntactically determines word order is in progress up to the EModE period would explain why the distribution of constituents within the VP complies with the syntactic principle of complements/arguments-first to a larger extent in EModE and LModE than in OE or ME.

4.1.2 Noun-and adjective-governed contexts, and 'complements-first'

In this section I will examine examples from the database of NPs and APs containing complements and adjuncts. As already pointed out in Section 3, in an attempt to retrieve uncontroversial instances of noun- and adjective-governed constructions, I have focused on the two prototypical types of complement clauses which rarely function as adjuncts, namely *that* and (*to* and bare) infinitive clauses. The exclusion of modifying infinitive clauses, such as that italicised in (31), from de Haan (1989:75), was carried out manually:

(31) the first thing *to do* is to send the paper

\(^{16}\) See Pintuz and Taylor (2006:260) and Los (2009:108; 2012:27) for support of the rigid subject-verb-object or subject-first syntax adopted after the ME period.
A corpus-based account of the placement of dependents in phrasal categories ... 519

It might be claimed that the findings of the VPs in Section 4.1.1, which contain clausal and non-clausal complements, cannot be compared to those of the NPs and APs in this section, with only clausal complements. However, in Pérez-Guerra (2016) I have focused on exclusively clausal complementation governed by verbs, adjectives and nouns and show that the statistical findings are very similar to those reported in this paper.

The hypothesis formulated in Section 4.1.1 that the principle of complements-first is quantitatively at work in the Modern periods, considerably so after the fixation of word order in the English clause, requires substantial data-based support. As shown in Table 1 above, in the case of the VPs, more than 60,000 examples were retrieved from OE to LModE. However, the examples of noun- and adjective-governed phrases containing complements and adjuncts do not abound in the database. The data are shown in Tables 2 and 3:

<table>
<thead>
<tr>
<th>Table 2, Noun-governed contexts (raw figures)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>OE</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>22</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3, Adjective-governed contexts (raw figures)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>OE</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

Figures 2 and 3 display the same information by relativising the proportions of complement-first and complement-last constructions. As already noted, the tendencies
seen in these figures must be interpreted very cautiously given that, especially in the OE period, the number of examples is very low:

Figure 2. Noun-governed contexts: Frequencies of complement–first (compl_first) and complement–last (compl_last) constructions

Figure 3. Adjective-governed contexts: Frequencies of complement–first (compl_first) and complement–last (compl_last) constructions

Figure 2 reveals that the percentage of complement-first NPs in the database is very much lower than the number of complement-first VPs. Whereas at least more
A corpus-based account of the placement of dependents in phrasal categories ... 521

than 70 percent of the VPs containing two postverbal dependents were complement-first in all periods from OE to LModE, the proportion of complement-first NPs does not reach even 20 percent in any period. As far as the statistical validation of the results is concerned, the increase in the number of complement-first examples from OE to ME is not significant (Wilson’s test [Poisson, assymetric]: p=0.08) and the data do not undergo any statistically recognisable trend from ME to EModE (p=1). Finally, the apparent increase of complement-first NPs seen in Figure 2 does not meet statistical verification either (p=0.191, $\chi^2=1.71$).

Summing up, the data do not show a clear statistical orientation towards a quantitative (and/or qualitative) change as regards the placement of complements and adjuncts in NPs. In fact, a chi-square test does not reveal the statistical significance of the data when applied to the complete distribution of NPs from OE to LModE (p=0.3195, $\chi^2=3.51$). I will come back to this conclusion in Section 5 when I discuss the different phrasal categories.

Figure 3 above shows the relative proportions of complement-first and complement-last APs in the different periods. Unlike the NPs in the database, we see here that the overall number of complement-first APs amounts to approximately 50 percent of examples. In other words, a continuum can be claimed to operate between VPs, APs and NPs as regards compliance with the syntactic principle of complements-first: 80 percent of verbs, 50 percent of adjectives and less than 20 percent of nouns are immediately followed by their complements. I will return to this in Section 5. However, the degree of the apparent variation seen in Figure 3 is not statistically significant, with the following chi-square tests: from OE to ME: p=0.3537, $\chi^2=0.86$; from ME to EModE: p=0.9203, $\chi^2=0.01$; from EModE to LModE: p=0.1738, $\chi^2=1.85$; for the whole distribution: p=0.2827, $\chi^2=3.81$. Even though the data are not statistically significant, the p value for the LModE examples, considerably lower than the same figures for the older periods, indicates not only that the proportion of complement-first APs is slightly higher in LModE, but also that the level of statistical significance of such a trend increases over time. I will come back to this in Section 5.
4.2 End-weight

In Section 2 I described the principle of end-weight, summarised as follows: “[p]hrases are presented in order of increasing weight” (Wasow 2002:3), characterising it as a valid performance-centred processing maxim which facilitates the planning, production and parsing of a given phrasal constituent. Hence, the so-called performance-related explanation leads to an assessment of the distance between the first and the second dependents, in order to determine whether the second dependents are longer or not. As Stowell (2006:239) maintains, “it has consistently proved to be virtually impossible to define ‘heaviness’ in a satisfactory way” and, consequently, there have been a number of proposals in the literature regarding the assessment of the degree of structural complexity of a construction (Stowell: 2006:239; see Wasow 1997 and Pérez-Guerra and Martínez-Insua 2010 for an overview). In this study I will evaluate the length of the dependents by measuring the average number of words.

Table 4 shows the average length of the constituents in my data. The information is also set out in Figures 4 (complement-first constructions) and 5 (complement-last constructions). Due to the low number of examples in the database, particularly NPs and APs, I will not approach the issue of compliance with end-weight from a diachronic perspective and will instead give only the mean lengths of the three phrasal categories. In the following section I will discuss the above findings in an attempt to determine the relative or absolute relevance of either the syntactic or the performance-driven explanations.

<table>
<thead>
<tr>
<th>Table 4. Word length of constituents in the phrase (mean values)</th>
</tr>
</thead>
<tbody>
<tr>
<td>complement-first</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>complement-last</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
A corpus-based account of the placement of dependents in phrasal categories ... 523

5. Results

In this section I will provide an interpretation for the findings outlined so far. First, I will measure the impact of the so-called syntactic principle of complements-first on
the design of VPs, NPs and APs with two post-head dependents. In Section 5.1 I will describe the data from a non-variationist perspective, looking again at the diachronic findings described in Section 4.1 and underlining the relevance of the Modern periods, LModE in particular, to the success of complements-first. In Section 5.2 I will then consider the so-called principle of end-weight. As already mentioned, the scarcity of examples in some of the phrasal categories justifies a synchronic treatment of the data.

5.1 Complements-first

According to the syntactic explanation, the ordering of the dependents in a phrase is determined by their syntactic status. In other words, the principle of complements-first rules that complements precede adjuncts.

Figure 6 displays the overall statistical tendencies of verb-, adjective- and noun-governed phrases and the proportions of combinations such as complement+adjunct (complement-first) and adjunct+complement (complement-last) in the database.

Figure 6 neatly shows the connection between the proportion of complement-first constructions and the syntactic category of the head. The data reveal that the more verbal the head (verb > adjective > noun) is, the more likely its compliance with
complements-first. More concretely, whereas the majority of the VPs are complement-first, almost 50 percent of the complements precede the adjuncts in the APs, and less than 20 percent of the NPs are complement-first.

On a theoretical basis, the prototypical status of verbs as structural governors has been argued from, for example, Stowell (2006), within the generative tradition and, from a cross-linguistic perspective, by Givón (1993:26-27), when he claimed that the morphological possibilities of a head should be taken as highlighting its potential as a governor since morphology implies syntactic integration, and this is most applicable to verbal heads. From less theoretical standpoints, frequency-based facts like the following give additional support to the archetypal status of verbs as governors or heads. First, their government potential, as evinced by the proportion of verbs with overt complements (the number of verb+complement constructions outweighs the number of nouns or adjectives followed by complements), backs up the characterisation of verbs as ‘better’ heads since they collocate with complements to a greater extent than, for example, nouns or adjectives. Second, the degree of paradigmatic versatility reflected by verbs as far as complement-taking patterns are concerned has a bearing on the special status of verbal heads in that verbs take part in a wider range of complementation patterns (monotransitive, ditransitive, copulative or intensive, adverbial, complex-transitive) than nouns or adjectives – to give some examples, nouns are not frequently found in copular relations with other dependents in NPs, and noun-governed constructions with two or more complements are very rare. Third, ellipsed nouns in NPs (eg. those Ø taking complements) are more common than ellipsed verbs in VPs; since heads cannot easily be dropped or deleted, it seems to be the case that verbs are (again) ‘better’ heads than nouns. Since frequency is taken here as a factor that leads to prototypicality, the above-mentioned facts highlight the archetypal status of verbs as far as headedness is concerned and explain why verb-governed constructions are subject to the principle of complements-first to a larger extent than, for example, NPs. Put differently, what I am arguing here is that the attested success of a syntactic principle such as complements-first in the VPs in the database finds an explanation in the prototypical status of verbs as syntactic heads or governors with respect to, for instance, complement-taking nouns. The examples which are adjective-governed occupy an intermediate position since adjectives are more verbal than nouns and thus occupy a middle point on the scale of government prototypicality, hence showing a higher
degree of compliance with syntactic principles such as complements-first than nouns.

Now, from a diachronic perspective, let us discuss the results of the historical analysis in Section 4.1. Figure 1 above showed that complements-first was in evidence for the majority of the VPs in the data. This scenario sheds light on the theoretical hypotheses described in the previous paragraphs that, first, the degree of compliance with the syntactic principle of complements-first is strongly influenced by the type of governing head and, second, that verbs are more prototypical governing categories than nouns or adjectives, which explains the successful outcome of complements-first in the VPs in the data. It should be noted that, as shown in Pérez-Guerra (2016), the type of complements investigated in this paper (NPs and clauses when governed by verbal heads, and strictly clauses in the APs and NPs) has not exerted an influence on the statistical results reported in this section. This amounts to saying that, although clauses are, on average, longer than NPs, end-weight has not distorted the overall proportions of complement-first VPs in the database.

With the focus on diachrony, Figure 1 above showed that the number of complement-first VPs has increased significantly over the course of time, and that the statistical consolidation of this trend took place in the Modern periods. This seems to bear out the proposal put forward in the current section, in that it underlines the connection between verbal status and observation of complements-first. In the absence of data for Present-Day English, EModE and LModE constitute the confirmation dot in the upward trend to the placement of complements immediately next to their governing heads.

Figures 2 and 3 above showed the slight increase of complement-first APs in the database and the modest decrease of complement-first NPs. Since neither of these was statistically validated, such tendencies must be treated with caution. The slight decreasing trend in the case of complement-first NPs in LModE might illustrate their separation from complements-first and this, according to my current argument, is a consequence of their non-verbal status. In turn, the mildly increasing bias towards higher numbers of complement-first APs in LModE seems to confirm the headedness hypothesis, in that adjectives are ‘more verbal’ than nouns. Further research, looking at Present-Day English, may substantiate the diachronic validity of the trends evinced by the LModE data and provide further confirmation of the headedness hypothesis.
5.2 End-weight

So far I have discussed the suitability of the syntactic explanation and have suggested that dependents are organised according to the complements-first principle in contexts of prototypical complementation, namely in VPs and, to a lesser extent, APs. In the examples containing heads which are not prototypical, that is, NPs, compliance with complements-first is less notable. In what follows I will focus on the so-called performance-related explanation. As already pointed out, performance-centred processing rules rely on the parsing cost of a given structure in order to determine its plausibility.

Figures 4 and 5 above shed light on the analysis of the performance account for the ordering of constituents within the phrases. The average length values reported in Table 4 above and sketched in Figures 4 and 5 are obviously biased by the fact that the complements governed by adjectives and nouns are always clausal, whereas the complements governed by verbs can be either clausal or phrasal. Given such a difference, the relevant information provided by Figures 4 and 5 does not lie in the average length values themselves but in the comparable length of the first and second dependents in the complement-first (Figure 4) and in the complement-last (Figure 5) examples. From such a comparative perspective, the figures show that in those cases where the syntactic principle of complements-first is at work the length of complements and adjuncts are more alike than in those instances which are not subject to complements-first. In detail, Figure 4 evinces that the adjuncts are, on average, less than three-word longer than the complements in complement-first VPs, and such a difference is even smaller when one compares complements and adjuncts in complement-first APs and NPs. By contrast, the second dependents (complements) are from 5- to 10-word longer than the first dependents (adjuncts) in the complement-last VPs, APs and NPs in Figure 5. In line with the scenarios above, processing rules such as end-weight (or Hawkins’ MiD) account for the distribution of dependents to a lesser extent in complement-first than in complement-last construction, and this is because the organisation of the dependents in the former constructions is not due to performance but rather to syntax. In other words, when the syntactic explanation fails and the construction is not complement-first, end-weight is crucial as regards the ordering of the complements and the adjuncts, and justifies the distribution of the dependents in the complement-last examples with
the three categories under analysis. This interpretation of the data leads to the conclusion that complements-first is prior to performance-based accounts.

Summing up, the discussion of the data here has suggested that the syntactic explanation, rather than the performance-related one, is more convincing as the major principle which rules the ordering of complements and adjuncts in the phrase. Such a conclusion is supported by the behaviour of particularly the noun-governed construction. It was noted above that the majority of the noun-governed examples do not conform to the syntactic explanation of complements-first. That said, Figure 4 illustrates that on those few occasions where the constructions headed by nouns are complement-first, the average length of the complements is close to the length of the adjuncts. This further argues for the conclusion that complements-first is a better fit than end-weight since even in the complement-first NPs the distribution of the dependents cannot be said to obey end-weight.

6. Summary and concluding remarks

This paper has reported on a corpus-based study of the ordering of dependents in noun, adjective and verb phrases. More specifically, it has focused on the combinations of complements and adjuncts in post-head position. The working hypothesis was that the distribution of adjuncts and complements in those phrases could be justified by syntactic principles such as complements-first and/or by processing rules such as end-weight. The data showed that complements-first is more powerful than end-weight, and that the role of the latter principle is subsidiary, especially when the former is not in evidence. These findings, then, do not corroborate Hawkins’ (1999:232) claim that “the biggest single predictor of relative orderings, for PPs as well as for other categories, is syntactic weight” (my italics). On the other hand, I have shown that compliance with complements-first is stronger when the construction contains a verbal or an adjectival head. This finding has been justified on the basis of the prototypical head status of verbs and, to a lesser extent, adjectives.

Since my data contained textual material from different periods (from OE to LModE) I was also able to confirm that the major result of this investigation, that is, the supremacy of the syntactic principle of complements-first as regards the
distribution of post-head dependents in VPs, NPs and APs, is especially operative after EModE onwards, that is, after the syntacticisation of word order in English, and is consolidated in LModE. On the one hand, the data neatly corroborate such a conclusion for VPs, that is, in the examples governed by the most prototypical heads, and less clearly so for the NP and the AP examples, in which statistical significance is not shown. On the other hand, the increase of complement-first phrases in the modern periods after the fixation of word order in the language gives supports to the claim that modern word order is ruled to a larger extent by syntax in English.

Issues such as the fine-grained characterisation of the different complementation strategies (e.g., that-clauses versus infinitive clauses), the effects of information-structure on the ordering of dependents, and the example-by-example correlation between complements-first and end-weight and its effects on the design of the constructions are left for further research.

**Corpora**


**References**


Los, Bettelou. 2012. The loss of verb-second and the switch from bounded to unbounded systems. In Anneli Meurman-Solin, María José López-Couso, and Bettelou Los (eds.),
A corpus-based account of the placement of dependents in phrasal categories ...

Pérez-Guerra, Javier. 2016 (In press). Do you investigate word order in detail or do you investigate in detail word order? On the connection between word order and headedness in the recent history of English. Corpus Linguistics and Linguistic Theory.
Pérez-Guerra, Javier and Ana Elina Martínez-Insua. 2010. Do some genres or text types become more complex than others? In Heidrun Dorgeloh and Anja Wanner (eds.), Syntactic variation and genre, 111-140. Berlin: Mouton de Gruyter.
patterns of nouns, adjectives and verbs over the last three centuries. Lanham: University Press of America.


