CrossMark

English Language and Linguistics 20.3: 463–486. © Cambridge University Press 2016
doi:10.1017/S1360674316000307

Testing two processing principles with respect to the extraction of elements out of complement clauses in English

GÜNTER ROHDENBURG
University of Paderborn
(Received 5 October 2015; revised 16 May 2016)

The present article contrasts two processing principles, the Domain Minimization Principle (e.g. Hawkins 1999, 2004) and the Complexity Principle (e.g. Rohdenburg 1996, 2007b) in structures involving the extraction of postverbal elements out of (competing) complement clauses. The Domain Minimization Principle may be described as a processing tendency which consists in minimizing the size and complexity of various domains including the filler–gap domain in cases like the following: *This is a problem (that) they had promised (that) they would tackle*. By contrast, the Complexity Principle represents a correlation between processing complexity and grammatical explicitness. It stipulates that more explicit (and typically more bulky) constructional options are favoured in cognitively more complex environments.

This article deals with two sets of rival complements. The behaviour of the first group of clausal alternatives is in line with Hawkins' prediction, though incompatible with the Complexity Principle. However, there is an even larger group of complement pairs whose distribution inside and outside extraction contexts is predicted by the Complexity Principle but unaccounted for by Domain Minimization. Thus, in extraction contexts, the visible effects of the two antagonistic principles are found with virtually complementary ranges of complement types. The article concludes by attempting to account for the kind of division of labour observed between the two principles under scrutiny. It will be suggested among other things that, in English, the marked infinitive (on its own or with an associated NP) enjoys a privileged or target status in extraction contexts.

1 Setting the scene

Over the last few decades, a growing number of linguists have investigated the role played by processing complexity in shaping grammatical variation and change. This article focuses on two relevant processing principles, the Domain Minimization Principle advanced by Hawkins (e.g. 1999, 2004) and the Complexity Principle invoked, for instance, in Rohdenburg (1996, 2007b), Vosberg (2006), Mondorf (2009) and Berlage (2014). The two principles profile entirely different perspectives and, so far, their application has generally involved complementary ranges of phenomena.

Three prominent areas in English analysed in terms of Domain Minimization include the variable placement of the particle with particle verbs (Lohse *et al.* 2004), the spread of gerundives to preverbal position at the expense of infinitives (Fanego 2010: 217–20), and the filler–gap domain in extractions, which is reviewed in this study. According to the Domain Minimization Principle, natural languages prefer

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

their processing domains to be as simple as possible both morphosyntactically and semantically.

By contrast, the Complexity Principle represents a correlation between two dimensions, processing complexity and grammatical explicitness, and it has usually been described as follows:

> In the case of more or less explicit constructional options, the more explicit one(s) tend to be preferred in cognitively more complex environments.

In other words, if the processing load goes up, so does the degree of grammatical explicitness. The principle covers a great variety of grammatical manifestations of cognitive complexity including those listed in (1).

(1) (a) discontinuous structures involving various kinds of insertions (e.g. Fanego 1990: 144–7; Rissanen 1991; Rohdenburg 1995, 1996, 2003; Vosberg 2006)
(b) voice contrasts (e.g. Rohdenburg 1996, 2002)
(c) number contrasts (e.g. Rohdenburg 2003)
(d) tense contrasts (e.g. Wasserman 1976; Rohdenburg 2003)
(e) subordinate clause negation (Horn 1978; Rohdenburg 2015)
(f) subject complexity (e.g. Quirk 1957; Wasserman 1976; Rohdenburg 2014)
(g) complex objects and other NPs (e.g. Rohdenburg & Schlüter 2000: 446–52; Berlage 2014)
(h) the contrast between finite and infinitival interrogative complements (Rohdenburg 2003)
(i) post-modification of relevant nouns (including other qualifying elements) (e.g. Rohdenburg 2002; Berlage 2014)
(j) the length of complement, relative and adverbial clauses (e.g. Quirk 1957: 108; Rohdenburg 1996: 164, 2006a: 55)
(k) gapping (Postal 1974: 129–31)
(l) right node raising (Postal 1974: 125–9)
(m) the rivalry between synthetic and analytic comparatives (Mondorf 2009)

Consider, for instance, the examples in (2), which illustrate constraint (1a). Compared with (2a), (2b) has been made more complex by inserting one of the four types of adverbs ending in *-ably, -ally, -antly* and *-ously* between the (intransitively used) verb *help* and the dependent infinitive.

(2) (a) These events helped (to) change her outlook.
(b) These events helped dramatically (to) change her outlook.

In keeping with the Complexity Principle, the structural discontinuity found in (2b) increases dramatically the proportion of the more explicit marked infinitive, as shown in table 1.

In this article, the two processing principles will be put to the test in examples like (3), which represent an infrequent and cognitively demanding kind of environment.

(3) (a) They had promised (that) they would tackle/to tackle this problem.
(b) This is a problem (which/that) they had promised (that) they would tackle Ø.
(c) This is a problem (which/that) they had promised to tackle Ø.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 1. *Marked and unmarked infinitives after the intransitively used active verb form* helped *in British newspapers*

		I *to*-inf.	II Ø	III total	IV % *to*-inf.
1	four different *-ly* adverbs inserted between *helped* and the infinitive (d91-00, g90-00)	25	6	31	80.6
2	other cases without insertions (samples from the first quarters of d95 and g95)	159	221	380	41.8

Example (3a) illustrates a canonical structure, where the subordinate verb (in *they would tackle*/*to tackle*) is followed by its direct object. In (3b, c), however, the direct object has been extracted from its canonical position by means of relativization, thus creating a gap (here indicated by Ø). In the relevant literature, the distance between the relative marker (*which*, *that* or zero) and the (canonical) site of the relativized element is usually described as a filler–gap domain. Examples such as (3b, c), where a postverbal element of the subordinate clause has been moved across the superordinate clause, contain unusually lengthy filler–gap domains and there is a general consensus that they are extremely complex structures to process (see e.g. Hawkins 1999, 2004).

In line with Domain Minimization, Hawkins (1999) demonstrates that English and various other languages display similar tendencies to minimize the size and complexity of this particular domain, and he proposes the following implicational scale for gaps in clause embeddings for English and, in fact, for all natural languages:

> It appears that infinitival phrases are most hospitable to gaps, that finite subordinate clauses are more resistant, while complex NP environments are most resistant of all. (Hawkins 1999: 263; see also Hawkins 2004: 193ff.)

In English, for instance, the contrast between complex NP environments and straightforward (finite) embedded clauses is evidenced in (4a, b) and that between finite and infinitival complements in examples like (5a, b).[1]

(4) (a) *Who did you hear the rumor [*sic*] that Mary saw? (Hawkins 1999: 267)
(b) Who did you hear that Mary saw? (ibid.)
(5) (a) *This is a task we do not know how we should deal with. (adapted from Hawkins 1999: 269; 2004: 198)
(b) This is a task we do not know how to deal with. (ibid.)

Note that the examples in (5) are even more complex than those in (3b–c): there are two gaps rather than just one. So (5a, b) certainly prove Hawkins' point. By contrast, the

[1] The rule excluding examples like (4a) was discovered by John Robert Ross (1967), who coined the phrase Complex NP Constraint to describe the phenomenon.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Complexity Principle would have the extraction contexts favouring the more explicit finite option. To that extent, it makes the wrong prediction in cases like (5). Moreover, the Complexity Principle, which focuses on possible variants in a given language rather than strictly excluded ones, has so far not concerned itself with contrasts like (4a, b). At this point it may be appropriate to reformulate the fundamental question to be posed with respect to the grammatical variants associated with (cognitively complex) extraction contexts: which of the variants is favoured in such environments, the morphosyntactically less complex one(s) (as stipulated by Domain Minimization) or the more explicit and typically more bulky one(s) (as predicted by the Complexity Principle)?

The database used for this study consists of several years of British and American newspapers (including *Time Magazine*) from the 1990s and early 2000s, the BNC, and a sizeable collection of historical data sets provided by Chadwyck-Healey and the Gutenberg project (see the bibliography for further details). Concerning the corpus analyses conducted, a few methodological remarks may be in order: While the focus is on British English (BrE), explorations concerning American English have usually been included only to increase the database or to draw attention to additional insights. In the case of rival complement options involving second and successive conjuncts, all analyses have been confined to the first in a series of conjuncts. The category of extractions generally disregards any examples of the fronting of direct speech and quoted thought. For reasons of both economy and homogeneity, the search procedures adopted have been restricted in a wide variety of ways. In addition, in terms of size and text type, the data sets have generally been selected to suit the purpose in question as far as possible.

Throughout, any detected duplicates and quotations from earlier periods than the (presumed) year of publication of the source in question have been excluded from consideration. The asterisk is used not only to refer to expressions found to be unacceptable but also to indicate the time span during which the authors concerned were born. Three levels of statistical significance will be distinguished: significant ($p < 0.05$ or $p < 0.025$), highly significant ($p < 0.01$) and extremely significant ($p < 0.001$). In the majority of sufficiently large samples, the figures compared reach significance levels ranging from $p < 0.01$ to $p < 0.001$. Therefore, explicit mention of the significance levels in question will usually be made only for tables containing what might appear to be doubtful cases.

The remainder of this article is organized as follows. Following up on examples (3)–(5), section 2 provides a brief survey of further acceptability-based findings. Section 3 goes on to analyse a novel range of complement pairs. Here, the argumentation will be supported throughout by distributional differences discovered in large electronically available databases (see the bibliography). In keeping with the Performance-Grammar-Correspondence Hypothesis (see Hawkins 2004: 3), the expectation is that the kinds of contrast predicted by the two processing principles should be reflected in text-distributional preferences. Section 4 provides a summary of major findings and some further discussion.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

2 A survey of previous acceptability-based findings

2.1 Finite clauses

The literature refers to three additional complement pairs involving finite complement clauses whose behaviour clearly supports Domain Minimization. They will be presented in turn.

Let us begin by taking a brief look at verbs complemented by *that*-clauses. On the basis of examples like (6a, b), Culicover (1999: 214–15) concludes that finite clauses with the complementizer *that* are more difficult to extract out of than those without.

(6) (a) *What book did Lee remark that you were reading. (Culicover 1999: 214)
(b) What book does Lee believe you were reading? (ibid.)

The conclusion is based on the following observation: while verbs such as *believe*, which potentially omit the complementizer, invariably permit the objects of their complement clauses to be extracted in cases like (6b), only some of the verbs that cannot drop the complementizer allow this procedure to be applied.

Next we turn to the so-called phenomenon of subject-to-object raising (or AcI) in contrasting examples like (7a, b), which have been discussed by a succession of linguists since Postal (1974).[2]

(7) (a) *They alleged Melvin to be a fraud.
(b) This is the man (that) they alleged to be a fraud.

Such examples demonstrate that – unlike their finite rivals in (8a, b) – the status of infinitival subject-to-object raising structures may be improved substantially in extraction contexts like (7b).

(8) (a) They alleged that Melvin was a fraud.
(b) This is the man (that) they alleged was a fraud.

Finally, this brings us to so-called *tough*-movement structures like (9a, b) and (10a, b) discussed amongst others by Berman (1973: 43), Ross (1974: 117) and Joseph (1980: 361).

(9) (a) *That ring is difficult to believe (that) Jack stole.
(b) That ring is difficult to believe Jack to have stolen.
(10) (a) *The car would be tough for me to suggest (that) we should leave at home.
(b) The car would be tough for me to suggest leaving at home.

Let us assume, as is commonly done in the literature on *tough*-movement (see e.g. Ross 1974: 17), that in cases like (9) and (10) the subject expressions *the ring* and *the car* have been extracted out of the complements of the verbs *believe* and *suggest*, respectively. On this assumption, these examples leave no doubt that – in line with Domain Minimization – non-finite complements, both infinitives and gerunds, are easier to extract out of than their finite counterparts.

[2] Other studies exploring the way extractions enhance the acceptability status of subject-to-object raising structures include Warner (1982: 134–57), Mair (1990, 1993) and Rohdenburg (2006c: 52–5).

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

2.2 *Gerunds*

It has occasionally been observed that – compared with contextually (near-)equivalent gerunds – specific infinitival complements may be strikingly improved in extraction contexts. Following Vosberg (e.g. 2006) and Fanego (e.g. 2007), I am assuming here that marked infinitives constitute more explicit grammatical structures than their gerundial counterparts. In addition, their morphosyntactic structure is, if anything, more complex than that of rivalling gerunds. The contrast between the two non-finite options will be discussed more fully in section 3.2 below.

In an article on comparatives, Pilch (1965: 51) observed the acceptability contrasts shown in (11a–c).

(11) (a) *I cannot stand him to be nasty.
(b) I cannot stand his being nasty.
(c) He is nastier than I can stand him to be.

It is evident that – unlike its gerundial counterpart – the acceptability of the object–infinitive construction, which may be treated as a manifestation of S-O Raising, is enhanced dramatically in extraction contexts like (11c). However, what Pilch failed to mention is the fact that the object–infinitive construction like many other examples of S-O Raising (see (7a, b), (8a, b) above or (14), (19), (20) and table 3 below), may also be improved by other kinds of extraction. Compare, for instance:

(12) (a) (?)They cannot stand this chap to be in charge.
(b) They cannot stand this chap being in charge.
(c) This is the chap they cannot stand to be/being in charge.

Similar observations have been made concerning examples like (13), which may involve unspecified object control with one or two directive verbs such as *advise* (Rohdenburg 2005).

(13) (a) (?)We advise to bring lifebelts along.
(b) We advise bringing lifebelts along.
(c) Lifebelts are among the items we advise to bring along.
(d) Lifebelts are among the items we advise bringing along.

Given that – in the cases surveyed – the infinitival structures are more explicit (and presumably also more complex in terms of grammatical nodes) than their gerundial counterparts, neither phenomenon can be handled in terms of Domain Minimization. They are, however, captured by the Complexity Principle.

2.3 *Infinitival clauses*

There are two further constructional rivals of subject-to-object raising structures featuring the string *to be*. In the first case, the sequence object + infinitive contrasts with its corresponding object predicative, its counterpart lacking *to be* as in (14) (see Quirk *et al.* 1972: 900).[3]

[3] The situation is paralleled in French by examples like (i)–(iii):

(i) Marie considère Robert intelligent. (Szamosi 1973: 652)
(ii) *Marie considère Robert être intelligent. (ibid.)
(iii) Les hommes que Marie considère être intelligent. (Szamosi 1973: 657)

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

(14) (a) ?We thought the boy to be honest.
(b) We thought the boy honest.
(c) The boy (that) we thought (to be) honest.

Similar contrasts are found in English with several object predicatives introduced by *as* and corresponding raising structures as in *regard x as y* vs *regard x to be y* or *represent x as y* vs *represent x to be y.*

In the second case, the raising structure is contrasted with the *for-to* construction as in (15).

(15) (a) He had asked for/*Ø the furniture to be removed. (Christian Mair p.c.; Rohdenburg 2005)
(b) Is that the furniture he had asked (for) to be removed? (ibid.)

With both types of contrast, the acceptability of the string object + *to be* is definitely enhanced in extraction contexts. In (14), the raising construction is both more complex morphosyntactically and more explicit than the competing object predicative. While the situation is incompatible with Domain Minimization, it is captured by the Complexity Principle. In (15), by contrast, it is only the shorter alternative without *for* that has dramatically improved its acceptability status in the extraction context. This behaviour fully conforms to Domain Minimization. Concerning the Complexity Principle, however, it would not be reasonable to argue that the (shorter) raising structure was any more explicit than its corresponding *for-to* construction. If anything, we have to assume the reverse to be true. As is well known, the complementizer *for* in (15) still preserves a lingering semantic and syntactic relationship with its prepositional source as in (16).

(16) He had asked for the removal of the furniture.

To that extent, the variant including *for* could be regarded as a more explicit option than its simpler counterpart. This means that (15b) contradicts the Complexity Principle: Crucially, it is the shorter and less explicit option whose acceptability is enhanced in the extraction context.

3 Corpus data

3.1 Finite clauses

Let us begin by revisiting two constraints mentioned earlier on. In the first case, we are comparing finite and infinitival complements of the verb *rejoice* as in (17) and (18) in a Late Modern English database.[4]

[4] To avoid the special effects of structural discontinuities analysed in depth by Vosberg (2006), in particular, examples containing any material between superordinate predicates and their complements have generally been discarded. In any case, examples featuring extractions and the remaining ones have always been treated the same in this respect.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 2. *Infinitival and finite* that*-clause complements after* rejoice *in Late Modern English corpora (ECF, NCF, EAF)*

		I *to*-inf.	II finite	III total	IV % *to*-inf.
1	extractions	30	–	30	100
2	remaining cases	529	153	682	77.6

Table 3. *S-O Raising* (… to be true) *and finite complements* (… is/are/was/were true) *after* believes/believed *in British newspapers (t90–04, g90–05, d91–00, d02, d04–05, i93–94, i02–05, m93–00)*

		I *to be true*	II finite	III total	IV % *to be true*
1	extractions	77	5	82	93.9
2	remaining cases	125	109	232	53.9

(17) (a) My daughter rejoiced that she had passed her exam.
(b) My daughter rejoiced to have passed her exam.
(18) (a) That is the exam my daughter rejoiced (that) she had passed.
(b) That is the exam my daughter rejoiced to have passed.

Since the implied subject of the infinitival complement is necessarily coreferential with the subject of the superordinate clause, the finite complements had to be similarly restricted. To be able to compare like with like, the analysis considers only those examples where the (explicit) subject of the finite *that*-clause complement also stands in a relation of coreference with the subject of the superordinate clause. The results of this comparison are found in table 2. Clearly, at $p < 0.01$, the corpus data corroborate the acceptability-based contrast seen in examples (5a, b) above: The marked infinitive is easier to extract out of than the finite clause.

The second type of contrast features finite clauses and subject-to-object raising structures associated with verbs like *believe* as in examples (19) and (20).

(19) (a) She still believes (that) the statement is true.
(b) She still believes the statement to be true.
(20) (a) This is the statement (that) she still believes is true.
(b) This is the statement (that) she still believes to be true.

Here the contrast seen in examples (7) and (8) above, which favours the sequence object + *to*-infinitive in extractions, is clearly confirmed (at $p < 0.001$) in table 3.

This brings us to a novel type of contrast, which involves directive verbs like *recommend* associated with the subjunctive and the rivalling modal structure represented by *should.* In order to make sure that we retrieve unambiguous and relatively common subjunctive predicates, we are comparing examples containing *be*-passives as in (21a, b).

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 4. should + *infinitive and the subjunctive in passive complements after* recommend *in British newspapers*

		I *should*	II subj.	III total	IV % *should*
1	extractions (t90–00, g90–00, d91–00, i93–94, m93-00)	26	6	32	81.3
2	remaining cases (t95–96, g95–96, d95–96, i93–94, m95–96)	103	146	249	41.4

Note: In order to compare like with like, any examples involving *no/not*-negation in the complement have been discarded, since they represent another complex environment (Rohdenburg 2015).

(21) (a) We recommended (that) they (should) be selected for the job.
(b) They were the people (that) we recommended (should) be selected for the job.

I take it that for a number of reasons set out in Rohdenburg (2015: 113–15) the subjunctive provides a (morphosyntactically) less complex and less explicit structure than its modal alternative. These arguments include the following:

1. Unlike the subjunctive, which shares its form with the imperative, the infinitive and finite verbs, the *should*-structure unambiguously represents a finite form.
2. The subjunctive behaves syntactically like a reduced form of the *should*-structure in at least two respects: *not*-negation does not trigger *do*-support, and the negator occupies the position it has in the corresponding modal alternative (see also Kjellmer 2009).

Turning now to the corpus analysis in table 4, we find that – contrary to Domain Minimization and in line with the Complexity Principle – it is the more complex and explicit modal variant that shows a special affinity for extraction contexts.

3.2 Gerunds

This section begins by contrasting gerunds and marked infinitives. The gerunds may be directly linked to their governing predicates as in (22) or they may involve various prepositions as in (23).

(22) (a) It was a scene which he had dreaded to see.
(b) It was a scene which he had dreaded seeing.

(23) (a) The projects that they were committed to support …
(b) The projects that they were committed to supporting …

In terms of grammatical complexity, the competing options in cases like (23) may seem to defy any straightforward ranking. However, a number of arguments presented by Vosberg (2003a, 2006) and Fanego (2007: 213–21) suggest that the marked infinitive

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

represents a more explicit clausal complement than the gerund. The arguments in support of this assessment include the following:

1. Unlike infinitives, gerunds stand in a paradigmatic relationship with (not necessarily synonymous) noun phrases, which underlines their lingering nominal orientation. Note furthermore that, in general, the choice of the gerund rather than the infinitive tends to correlate with a given verb's affinity for nominal complements (see e.g. Kjellmer 1980; De Smet 2010, 2012: 219, 254ff.).
2. The cue validity of gerunds is comparatively low, in particular in extraction contexts where they occur in clause-final position (concerning the notion of cue validity, see e.g. MacWhinney *et al.* 1985). Unlike the marked infinitive, which immediately signals its verbal function, the recognition of the gerund as a clausal complement is more highly dependent on (the adjacency of) a governing verb and its characteristic nominal complements and adjuncts.
3. Previous explorations have shown that – compared with all kinds of gerunds – the marked infinitive displays a special affinity for some of the complexity factors listed in (1) above**,** including the use of insertions (see, in particular, Vosberg 2006) and complement negation (Rohdenburg 2015). Crucially, these complications favour the choice of several clearly more explicit types of subordinate clause over less explicit ones, replacing, for instance, the marked infinitive by the *that*-clause or finite clauses involving the subjunctive by those containing *should* + infinitive (as shown in the preceding section).

Thus it is reasonable to assume that the marked infinitive represents a more explicit clausal structure than the gerund. In line with the Complexity Principle, Vosberg (2003a, b, 2006) and Rohdenburg (2006b: 153–5, 2007a: 197–9, 2007b: 228–31) have shown in various corpus analyses that all kinds of gerunds are more difficult to extract out of than corresponding marked infinitives. Additional evidence supporting these findings is presented in table 5, dealing with directly linked gerunds after the verb *dread* as in examples (22a, b) above and tables 6 and 7, which present a selection of cases involving a prepositional linkage. Let us now consider these analyses in some more detail.

Like many other verbs, *dread* has generally replaced infinitival complements with gerunds. Table 5 shows that, in straightforward environments, the infinitive is used today in a measly 3 per cent of all cases. By contrast, in extraction contexts, it still accounts for over a quarter of all examples. At $p < 0.001$, the figures in rows A1 and A2 are extremely significant. Following a lead given by Bolinger (1979), the 100 extraction cases have been subjected to a closer analysis. Here, a distinction is drawn between non-finite verbs occurring in sentence-final position and elsewhere. We can see at a glance that the advance of the gerund has been delayed to a strikingly greater extent in sentence-final position than in all other environments. This confirms Bolinger's assumption that sentence-final position is least compatible with gerunds, but it also suggests that the distribution of the gerund and the infinitive in extractions might to some extent be motivated by other factors than processing complexity.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 5. *Marked infinitives and directly linked gerunds after* dread *in British newspapers (t90–01, g90–00, d91–00, m93–00) and the BNC)*[5]

		I *to*-inf.	II *-ing*	III total	IV % *to*-inf.
A	all examples				
	1 extractions	27	73	100	27
	2 remaining cases	30	902	932	3.2
B	all extraction cases (=A1)				
	1 [sentence-final] = immediately preceding /./;/,/:/?/-/	18	23	41	43.9
	2 remaining cases	9	50	59	15.3

Crucially, however, even the occurrence of extraction gaps in other than sentence-final position (=B2) provides an extremely significant contrast with the remaining canonical examples (=A2) (at $p < 0.001$).

The rivalry between the marked infinitives and prepositional gerunds is presented in tables 6 and 7. The predicates dealt with in table 6 are illustrated in examples (24)–(27). They involve two verb–noun collocations and two adjectives.

(24) (a) They have a very good reason to ignore/for ignoring that event.
(b) The event that we have a very good reason to ignore/for ignoring …

(25) (a) Some of us had the privilege to experience/of experiencing that event.
(b) The event some of us had the privilege to experience/of experiencing …

(26) (a) She was proud to have seen/of having seen that match.
(b) The match that she was proud to have seen/of having seen …

(27) (a) He was disposed to support/to supporting the project.
(b) The project he was disposed to support/to supporting/toward(s) supporting.

Again, the analyses distinguish between two kinds of contrasting structures, extraction contexts and all remaining environments. Paralleling the findings of e.g. Vosberg (2003a, 2006) and Rohdenburg (2007a, b), the results leave no doubt that in every single case the marked infinitive is proportionally more common in extractions than elsewhere. While the figures for the contrasting environments are extremely significant in A–C, those in D reach only the basic level of significance at $p < 0.025$.

This brings us to table 7, which deals only with the adjective *accustomed* as illustrated in (28).

[5] The evidence presented in this table is restricted to examples of gerunds and infinitives occurring within two words to the right of the verb *dread*. The analysis disregards a range of (knock-out) contexts which are confined to infinitival complements. They include the following:

(i) interrogative complements associated with subordinate verbs
(ii) the use of the subordinate verb form *think*
(iii) direct objects associated with the subordinate verb *imagine*

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 6. *Marked infinitives and prepositional gerunds after two verb–object collocations and two adjectives (used non-attributively) in British databases*

		I *to*-inf.	II *-ing*	III total	IV % infinitive
A	*have (…) reason(s)* (t90–00, BNC)				
	1 extractions	89	1	90	98.9
	2 remaining cases	4,066	760	4,826	84.3
B	*have the (…) privilege* (t90–03, g90–00, d91–00)				
	1 extractions	28	43	71	39.4
	2 remaining cases	77	753	830	9.3
C	*proud* (m93-00)				
	1 extractions	123	2	125	98.4
	2 remaining cases	1,199	154	1,353	88.6
D	*disposed* (t90-04, g90-04, d91-00, i93-94, i02-04, m93-00)				
	1 extractions	73	1	74	98.6
	2 remaining cases	975	115	1,090	89.4

Note: Examples involving adjectives premodified by *too* as in *too proud*, which invariably select the infinitive, have been excluded from consideration.

Table 7. *Marked infinitives and prepositional gerunds after* accustomed *in British newspapers (d91–98, m93–00), the BNC, LOB and FLOB*

		I *to*-infin.	II *-ing*	III total	IV % *to*-inf.
A	all examples				
	1 extractions	58	65	123	47.2
	2 remaining cases	11	1,174	1,285	8.6
B	all extraction cases (=A1)				
	1 monosyllabic verbs	41	56	97	42.3
	2 more complex verbs	17	9	26	65.4

(28) (a) We are accustomed to make/making these distinctions.
(b) The distinctions we are accustomed to make/making …

We know that after *accustomed* the infinitive has been largely replaced by the gerund (see e.g. Rudanko 1999: 9–10, 2000: 90–1, 2010; Vosberg 2006: 238–44; Rohdenburg 2006b: 153–5). Not surprisingly, the data in rows A1–2 of table 7 show that the replacement process is less far advanced in extraction contexts, thus confirming

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

previous conclusions. Consider now in rows B1–2 the more detailed analysis carried out on the 123 extraction cases. According to Zipf (1935), the length and complexity of a given expression is in general positively correlated with its semantic specificity and negatively correlated with its frequency of occurrence. In line with Zipf's prediction and Vosberg's in depth analysis, the data in rows B1–2 reveal that the replacement process has been delayed, in particular, with those subordinate verbs that are more complex phonologically and (most likely) also semantically and which can be assumed to be difficult to process. The figures in B1–2 are significant at $p < 0.025$.

We turn next to the contextually equivalent types of constructions in (29) and (30), which will be referred to as *help -ing* and *help but*.

(29) (a) They could not help criticizing the project.
(b) They could not help but criticize the project.

(30) (a) This is a project (that) they could not help criticizing.
(b) This is a project (that) they could not help but criticize.

I take it that – for a number of reasons put forward in Rohdenburg (2013) – *help but* represents a more explicit structure than *help -ing*. In addition to those associated with the nominal character of gerunds mentioned above, they include the following:

1. Corpus analyses have shown that *help but* is attracted to several secondary, derived or more complex clause types including the use of inanimate subjects of the superordinate predicate and the use of the passive and/or perfect of the dependent verb phrase.
2. Unlike *help -ing*, *help but* is invariably associated with a verbal item, which selects the same range of complements or adjuncts as other verb forms. In addition, the verb stem after *help but* alternates occasionally with the marked infinitive, which confirms its status as a bare infinitive.
3. Moreover, the use of *but* in the sense of 'except' reinforces or clarifies the semantic orientation of the complex *help* + dependent clause. The situation is paralleled by that of *prevent s.o. (from) -ing* in British English, where the optional use of *from* helps to clarify the semantics of the overall construction.

On the basis of the evidence in table 8, we can assume that *help but*, which we regard as more explicit than *help -ing*, has always been more strongly attracted to extraction contexts than its gerundial rival. The three analyses reach the following levels of statistical significance: $p < 0.025$ in A, $p < 0.01$ in B and $p < 0.05$ in C. In any case, since the filler–gap domain of *help -ing* is shorter than that of *help but*, both the historical and present-day data fail to accord with the prediction of Domain Minimization.

The following pair of complements involves two different kinds of contextually equivalent gerunds: perfective and non-perfective gerunds after retrospective verbs. Here, simple (non-perfective) and complex perfective gerunds as in (31a, b) have ousted the earlier perfective infinitive as in (31c).

(31) (a) I remember/recollect meeting her sister before.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 8. *The rivalry between* help but *and* help -ing *in British newspapers of the early 1990s and early 2000s and a series of American fictional databases of the eighteenth and nineteenth centuries: t90–91, g90–91, d91–92, d94–95, m94–95; t03–04, g04–05, d04–05; AD (*1750–*1849), EAF (*1750–*1827), LNC/A (*1860–*1869), ETC/A (*1870–*1894)*

		I *help but*	II *help -ing*	III total	IV % *help but*
A	BrE: early 1990s				
	1 extractions	14	15	29	48.3
	2 remaining cases	520	1,243	1,763	29.5
B	BrE: early 2000s				
	1 extractions	34	19	53	64.2
	2 remaining cases	1,390	1,580	2,970	46.8
C	AmE: late 18th to early 20th centuries				
	1 extractions	11	44	55	20
	2 remaining cases	139	1,415	1,554	8.9

(b) I remember/recollect having met her sister before.
(c) *I remember/recollect to have met her sister before.

Vosberg (2003a: 308–11, 2006) has shown that – while it was current – the perfective infinitive used to be preferred in extraction contexts over both types of gerunds. Whereas this observation is incompatible with Domain Minimization, it is accounted for by the Complexity Principle. Concerning the two types of gerunds themselves, it is obvious that, with retrospective verbs and backward looking gerunds, the perfective gerund should be regarded as the more explicit of the two. Now Vosberg (2003a: 311–12, 2006: 180, 201) finds that with both *remember* and *recollect* it is the perfective variant that seems to have always displayed a special affinity for extraction contexts as in (32). Again, Domain Minimization fails to account for the fact that the shorter variant is, relatively speaking, less common in such environments.

(32) This is the girl I remember/recollect having met before.

In the meantime, I have conducted several companion analyses involving *admit*, *recall* and *remember* in a large newspaper collection from the 1990s and the early 2000s. Even if the proportion of the perfective variant is very low, the contrast between the two rival options seen in Vosberg's (2006) mainly historical work still remains visible in large newspaper databases. Compare, for example, the evidence for *recall* in British English in table 9 and that for *remember* in table 10 for British and American English. With both verbs, and unlike Vosberg's more comprehensive analyses, the category of extractions has been confined to the relativization of direct objects as illustrated for the perfective gerund in example (32) above. Accordingly, the category of the remaining

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 9. *Perfective and non-perfective gerunds after* recall *in British newspapers (t90–00, g90–00, d91–00, m93–00)*

		I perfective	II non-perfective	III total	IV % perfective
1	extractions	16	56	72	22.2
2	remaining cases	109	1,916	2,025	5.4

Note: The database for the non-perfective gerunds in column II comprises the following samples: the first quarters from t93, g93, d03, m93, t98, g98, m98, and the first and second quarters from d98.

Table 10. *Perfective and non-perfective gerunds after* remember *in British and American newspapers*

		I perfective	II non-perf.	III total	IV % perfective
A	BrE (t90–01, g90–00, d91–00, m93–00)				
	1 extractions	16	390	406	3.9
	2 remaining cases	115	8,454	8,569	1.3
B	AmE (L92–99, D92–95, W90–92, N01)				
	1 extractions	11	204	215	5.1
	2 remaining cases	31	3,974	4,005	0.8

Note: The databases analysed for the non-perfective gerunds are as follows. BrE: first quarters of t95, g95, d95 and m95; AmE: L99, D95, W91 (excluding the supplement *Insight on the News*) and the first two quarters of N01.

gerunds not involving extractions is also represented by verbs governing direct objects as shown in (31a, b). Furthermore, the totals for the non-perfective gerunds have been estimated on the basis of the attested examples in representative samples of the overall databases used. Due to the large number of examples involved, the figures in all three analyses are extremely significant at $p < 0.001$, thus supporting the earlier observations made by Vosberg (2003a, 2006).

3.3 Infinitival clauses

There are two remaining pairs of complements contrasting marked infinitives with related structures. The first pair contrasts passive raising structures after the directive predicates *ordered* and *would like* as in (33) with their counterparts lacking *to be*.

(33) (a) We ordered/would like the horse (to be) dope tested.
(b) This is the horse we ordered/would like (to be) dope tested.

Intriguingly, examples like (33) behave very differently from those in (14) above. With both predicates in (33), the frequency of the shorter and less explicit variant involving

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 11. *Passive raising structures after* ordered *and* would like *and their reduced variants lacking* to be *in British and American newspapers (*ordered *in BrE: t90–00, g90–00, d91–00, i93–94;* would like *in AmE: L92–94, D92–95, W90–92, N01)*

		I *to be -ed* etc.	II Ø	III total	IV % *to be -ed* etc.
A	BrE: *ordered*				
	1 extractions	21	15	36	58.3
	2 remaining cases	59	6	65	90.8
B	AmE: *would like*				
	1 extractions	3	33	36	8.3
	2 remaining cases	47	66	113	41.6

to be-deletion is clearly increased in extraction contexts over its more complex and explicit counterpart. Compare the evidence in table 11, which – running counter to the Complexity Principle – conforms to the prediction of Domain Minimization.

The second pair of complements compares marked and unmarked infinitives as illustrated in (34).

(34) (a) He helped/had helped (to) establish the system.
(b) This is the system he helped/had helped (to) establish.

The more explicit type including the infinitive marker is at present being replaced by the less explicit variant, with American English spearheading the change. As has been shown in a number of studies, several of the complexity factors mentioned in list (1) above, such as object complexity or complement negation (see Berlage 2014: 196–230; Rohdenburg 2015: 111), tend to slow down the erosion of the infinitive marker in both transitive and intransitive uses of *help.* However, for the sake of convenience, this article is exclusively concerned with intransitively used *help* governing transitive verbs like *establish* as shown in examples (34a–b). Concerning extractions as in (34b), Domain Minimization predicts that the unmarked infinitive should be favoured over the more complex marked infinitive.

Consider now the large-scale investigation summarized in table 12, which deals with two superordinate phrases, *he helped* and *he had helped* in the *Los Angeles Times* and the *Guardian*. Allowing for the fact that American English appears to be much further advanced in this area than British English, it is seen that both varieties display the same general tendency, which is diametrically opposed to Domain Minimization.[6] Throughout, and in line with the Complexity Principle, the more explicit marked infinitive is proportionally used more frequently in extraction contexts than elsewhere.

[6] Compared with the past tense, the pluperfect has delayed the erosion of the infinitive marker in both British and American English, a phenomenon that would certainly repay closer analysis.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 12. *Marked and unmarked infinitives of transitive verbs after intransitive* help *in British and American newspapers*

		I *to*	II Ø	III total	IV % *to*
A	BrE: *he helped* (g90–03)				
	1 extractions	187	229	416	45.0
	2 remaining cases	404	863	1,267	31.9
B	BrE: *he had helped* (g90–03)				
	1 extractions	72	52	124	58.1
	2 remaining cases	54	70	124	43.5
C	AmE: *he helped* (L92–96)				
	1 extractions	50	441	491	10.2
	2 remaining cases	64	1,082	1,146	5.6
D	AmE: *he had helped* (L92–99)				
	1 extractions	11	30	41	26.8
	2 remaining cases	7	63	70	10.0

While the figures in analyses A and C are extremely significant at $p < 0.001$, those for B and D only reach the basic level of significance at $p < 0.025$.

4 Conclusion

This article has surveyed – in both canonical environments and non-canonical extraction contexts – a total of sixteen complement pairs (including one or two variants). The overview in table 13 identifies seven cases that are handled only by the Domain Minimization Principle and nine others that are only accounted for by the Complexity Principle.

Thus the visible effects of the two principles are found in virtually non-overlapping ranges of application. It would follow that we cannot dispense with the basic insights afforded by either principle.

For the sake of clarity, the sixteen contrasting pairs in table 13 have been subdivided into three groups, A, B, and C. Group A is confined to finite complements, which are paired with (a) three other finites (A1, A2, and A6) and (b) three non-finites (A3, A4 and A5). Group B contrasts the remaining gerunds with infinitival constructions. Group C contains the remaining infinitival constructions, which are paired with (a) other infinitival structures and (b) reduced structures involving *to-be* deletion.

Since, strictly speaking, A1 does not feature two constructional variants, it falls outside the domain dealt with by the Complexity Principle. This is why it will also be excluded from further consideration. Of the remaining five pairs of complements

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Table 13. *Overview of competing complement clauses in extraction contexts*

		I Domain Minimization		II Complexity Principle	
A	finite complements				
	1 finite S within S in complex NP	+		−	
	2 presence or absence of *that*	+		−	
	3 finite vs infinitival clauses	+	(*to*-inf.)	−	
	4 finite S vs gerund	+		−	
	5 finite vs S-O raising structures	+	(NP + *to*-inf.)	−	
	6 *should* + infinitive vs subjunctive	−		+	(*should* + inf.)
B	gerunds (other than A4)				
	1 NP('s) + *-ing* vs S-O Raising	−		+	(NP + *to*-inf.)
	2 *advise* + *-ing* vs *advise* + *to*-infinitive	−		+	(*to*-inf.)
	3 gerund (directly linked) vs *to*-infinitive	_		+	(*to*-inf.)
	4 prepositional gerund vs *to*-infinitive	_		+	(*to*-inf.)
	5 *help -ing* vs *help but* + bare infinitive	_		+	(*but* + bare inf.)
	6 perfective vs non-perfective gerunds	_		+	
C	(remaining) infinitival clauses				
	1 S-O Raising after *think* etc. vs counterparts lacking *to be*	_		+	(NP + *to*-inf.)
	2 passive raising structures after *order* etc. vs counterparts lacking *to be*	+	(NP + *-ed*)	_	
	3 *for-to* structures vs S-O Raising after *ask*	+	(NP + *to*-fin.)	_	
	4 marked vs unmarked infinitives after *help*	_		+	(*to*-inf.)

in group A, the first four run counter to the predictions of the Complexity Principle. They do, however, behave as predicted by Domain Minimization. Notice that in A2 one of these pairs features finite complements with or without the complementizer *that.* By contrast, A6, the only other pair consisting of two finite complements, shows a behaviour that is only captured by the Complexity Principle. This leaves us wondering what motivates the different behaviour of A2 and A6. Could it be that – apart from increasing the length of the filler–gap domain – the complementizer plays a decisive role in clearly separating the complement from the superordinate clause?

Concerning five of the gerunds in B paired with infinitival structures, we can say that the less explicit gerund is outranked in extraction contexts by two more explicit and morphosyntactically more complex types, (a) the marked infinitive (preceded by an NP in B1) and (b) the combination consisting of *but* + bare infinitive. Furthermore, in the pair opposing two different gerunds (B6), it is the more formally complex and explicit perfective gerund that is favoured over the much more common non-perfective type.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Turning now to group C, we find that, in each subgroup mentioned above, one of the contrasts is only accounted for by the Complexity Principle while the other conforms only to Domain Minimization. Again we are left to speculate as to what motivates the contrasting behaviour of C1–2 and C3–4. Could it be, for instance, that – in parallel with *that* in A2 – the demarcating function of the complementizer *for* is responsible for disfavouring the extraction context? While the reduced constructions with verbs like *think* in C1 are typically confined to stative and non-future predicates, those of verbs like *order* in C2 tend to involve non-stative and future predicates. Moreover, apart from the occasional use of adverbs like *out*, reduced structures after *order* are typically confined to passive uses of the perfect participle. Thus the clear-cut preference for *to be*-deletion in passive raising structures after directive predicates might be partially explained by their unusually high cue validity.

We have seen, then, that the marked infinitive, on its own or with an associated NP in raising structures, enjoys a privileged or target status in extraction contexts. Conforming, as it were, to Domain Minimization, it outranks finite clauses and *for-to* structures, but conforming, as it were, to the Complexity Principle, it also outranks all kinds of gerunds, the unmarked infinitive (after *help*) and object predicatives or reduced raising structures involving *to be*-deletion – in C1 – after verbs like *think*. The special role played by the marked infinitive may be due to the fact that it represents a good compromise in a hypothetical trade-off between Domain Minimization and the Complexity Principle: it is neither a very complex morphosyntactic structure nor an overly inexplicit one. For instance, with the prominent subject-to-object raising verbs, the verbs of knowing, thinking and saying, the marked infinitive in S-O raising structures is preferred over both the finite clause, which is structurally too complex, and the object predicative (produced by *to be*-deletion), which is less explicit.

Let us briefly return to those pairs of complements where – in extraction contexts – Domain Minimization wins out over the Complexity Principle. Here, it should be pointed out that with other complexities (such as discontinuities, complement negation, and passivization) we can expect the usual alignment of the more complex environment with the more explicit variant to reassert itself. This is the case, for instance, with the following pairs of complements:

– *that*-clauses with or without the complementizer (e.g. Wasserman 1976; Fanego 1990: 144–7; Rissanen 1991).
– *that*-clauses contrasting with marked infinitives or gerunds (e.g. Vosberg 2006; Rohdenburg 2015).
– *for-to* structures contrasting with subject-to-object raising ones (e.g. Postal 1974).

Furthermore, as is suggested by the analysis of *accustomed* in section 3.2, the processing of the filler–gap domain is not just a syntactic issue. According to Hawkins (1999: 271ff), the human processor also prefers to minimize the semantic information that needs to be processed in a filler–gap domain. However, what still has to be explained is why this tendency should affect gerunds so much more strongly than

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

marked infinitives. And we also need to account for the fact – shown in the case of the verb *dread* in section 3.2 – that gerunds are most strongly objected to in sentence-final position. Could it be, as suggested by Bolinger (1979), that prosodic prominence in final position has the effect of further weakening an already weak structural choice?

In addition, there are still a few contrasts left in the area of clausal complementation that have yet to be sorted out. They include prepositional vs directly linked gerunds as in (35) and marked infinitives vs pseudo-coordinated structures as in (36).

(35) (a) They had difficulty (in) coming to grips with that problem.
(b) Lack of time prevented him (from) finishing the job.
(36) You should try to/and meet the deadline.

In (35), the prepositional gerunds clearly represent the more explicit grammatical alternatives. However, the situation in (36) is less straightforward. While the rival options appear to be equally complex in terms of grammatical nodes, the marked infinitive has been shown to constitute the more explicit structure (see Rohdenburg 2015: 112–13). In both cases, then, it is the more explicit rivals that – according to the Complexity Principle – should be favoured in relevant extraction contexts. By contrast, Domain Minimization could be expected to favour extractions out of directly linked gerunds over prepositional ones in (35), while having to leave open the case of extractions out of the complements in (36). Finally, going beyond the area of complement clauses, we may encounter similar contrasts with extractions out of more or less complex and explicit purpose clauses such as (37a, b).

(37) (a) ?He is the sort of player most clubs go out of their way in order/so as to attract.
(b) He is the sort of player most clubs go out of their way to attract.

The rivalry between cases like these would appear to be captured by Domain Minimization rather than the Complexity Principle.

Author's address:

Department of English and American Studies
University of Paderborn
Warburger Strasse 100
33098 Paderborn
Germany
rohdenburg@onlinehome.de

Electronic databases

AD *American Drama.* 2005. Enigma Corporation Inc./Proust Information and Learning Company. Ann Arbor, MI/Cambridge, UK (size of the reduced version used: 16,849,683 words)

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

BNC	*British National Corpus* 1995. version 1.0. BNC Consortium/Oxford University Computing Services. (100,000,000 words)
d91–00, 02, 04–05	*Daily Telegraph and Sunday Telegraph* on CD-ROM 1991–2000, 2002, 2004–5. Chadwyck-Healey/ProQuest. (478,837,273 words)
D92–95	*Detroit Free Press* on CD-ROM 1992–1995. Knight Ridder Information Inc. (102,989,512 words)
EAF	*Early American Fiction* 2000. Authors born between 1744 and 1827. Chadwyck-Healey. (34,634,666 words)
ECF	*Eighteenth-Century Fiction* 1996. Authors born between 1660 and 1752. Chadwyck-Healey. (9,702,696 words, omitting duplicates)
ETC	*Early Twentieth Century Corpus* – a selection of British and American writings by authors born between 1870 and 1894. Source: Project Gutenberg. Compiled in the Research Project 'Determinants of Grammatical Variation in English', University of Paderborn. (16,351,681 words)
ETC/A	American writings in the ETC. (11,550,273 words)
FLOB	Match of LOB compiled at Freiburg University (representing written British English from 1991). ICAME. (1,000,000 words)
g90–05	*Guardian* (including *The Observer* 1994–2005) on CD-ROM 1990–2005. Chadwyck-Healey/ProQuest. (645,817,821 words)
i93–94, 02–05	*Independent and Independent on Sunday* on CD-ROM 1993–4, 2002–5. ProQuest. (242,608,117 words)
L92–95	*Los Angeles Times* on CD-ROM 1992–5. Knight Ridder Information Inc. (320,016,164 words)
L96–99	*Los Angeles Times* 1996–9 (courtesy of The Los Angeles Times Editorial Library). (275,506,490 words)
LNC	*Late-Nineteenth-Century Corpus* – a selection of British and American writings (complementary to the EAF and NCF) by authors born between 1830 and 1869. Source: Project Gutenberg. Compiled in the Research Project 'Determinants of Grammatical Variation in English', University of Paderborn. (47,677,728 words)
LNC/A	American writings in the LNC. (26,859,926 words)
LOB	*The Lancaster–Oslo/Bergen Corpus* on CD-ROM 1961. ICAME. (1,000,000 words)
m93–00	*Daily Mail and Mail on Sunday* on CD-ROM 1993–2000. Chadwyck-Healey. (206,762,410 words)
N01	*New York Times* on CD-ROM 2001. ProQuest. (52,132,97 words)
NCF	*Nineteenth-Century Fiction* 1999–2000. Chadwyck-Healey. (37,589,837 words)
t90–04	*The Times and The Sunday Times* on CD-ROM 1990–2004. Chadwyck-Healey/ProQuest. (729,848,339 words)
TAL89–94	*Time Magazine* on CD-ROM 1989–1994. (12,123,886 words)
W90–92	*Washington Times* (including *Insight on the News* 1990–1992) on CD-ROM 1990–2. Wayzata Technology. (93,889,488 words)

References

Berlage, Eva. 2014. *Noun phrase complexity in English*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Berman, Arlene. 1973. A constraint on tough-movement. *Chicago Linguistic Society* (*CLS*) 9, 34–43.
Bolinger, Dwight. 1979. The jingle theory of double *-ing*. In David J. Allerton, Edward Carney & David Holdcroft (eds.), *Function and context in linguistic analysis: A festschrift for William Haas*, 41–56. Cambridge: Cambridge University Press.
Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory and language acquisition*. Oxford: Oxford University Press.
De Smet, Hendrik. 2010. English *-ing* clauses and their problems: The structure of grammatical categories. *Linguistics* 48(6), 1153–93.
De Smet, Hendrik. 2012. *Spreading patterns: Diffusional change in the English system of complementation*. Oxford: Oxford University Press.
Fanego, Teresa. 1990. Finite complement clauses in Shakespeare's English, part 2. *Studia Neophilologica* 62(2), 129–49.
Fanego, Teresa. 2007. Drift and the development of sentential complements in British and American English from 1700 to the present day. In Javier Pérez-Guerra, Dolores González-Alvarez, Jorge L. Bueno-Alonso & Esperanca Rama-Martínez (eds.), *'Of varying language and opposing creed': New insights into Late Modern English*, 161–235. Bern: Peter Lang.
Fanego, Teresa. 2010. Variation in sentential complements in eighteenth- and nineteenth-century English: A processing-based explanation. In Raymond Hickey (ed.), *Eighteenth-century English: Ideology and change*, 200–20. Cambridge: Cambridge University Press.
Hawkins, John A. 1999. Processing complexity and filler–gap dependencies across grammars. *Language* 75, 244–85.
Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
Horn, Laurence R. 1978. Some aspects of negation. In Joseph Greenberg (ed.), *Universals of Language*, vol. IV, 127–210. Stanford, CA: Stanford University Press.
Joseph, Brian. 1980. Linguistic universals and syntactic change. *Language* 56, 345–70.
Kjellmer, Göran. 1980. Accustomed to swim: Accustomed to swimming. On verbal forms after *to*. In Jens Allwood & Magnus Ljung (eds.), *ALVAR: A linguistically varied assortment of readings. Studies presented to Alvar Ellegård on the occasion of his 60th birthday* (= Stockholm Papers in English Language and Literature 1), 75–99. Stockholm: University of Stockholm.
Kjellmer, Göran. 2009. The revived subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English*, 246–56. Cambridge: Cambridge University Press.
Lohse, Barbara, John A. Hawkins & Thomas Wasow. 2004. Processing domains in English verb–particle constructions. *Language* 80(2), 238–61.
MacWhinney, Brian, Csaba Pléh & Elizabeth Bates. 1985. The development of sentence interpretation in Hungarian. *Cognitive Psychology* 17, 178–209.
Mair, Christian. 1990. *Infinitival complement clauses in English: A study of syntax in discourse*. Cambridge: Cambridge University Press.
Mair, Christian. 1993. A cross-linguistic functional constraint on *believe*-type raising in English and selected other European languages. *Papers and Studies in Contrastive Linguistics* 28, 5–19.
Mondorf, Britta. 2009. *More support for more-support: The role of processing constraints on the choice between synthetic and analytic comparative forms*. Amsterdam: John Benjamins.
Pilch, Herbert. 1965. Comparative constructions in English. *Language* 41(1), 37–58.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Postal, Paul. 1974. *On raising*. Cambridge, MA: MIT Press.
Quirk, Randolph. 1957. Relative clauses in educated spoken English. *English Studies* 38, 97–109.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1972. *A grammar of contemporary English*. London: Longman.
Rissanen, Matti. 1991. On the history of *that/zero* as object clause links in English. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 272–89. London: Longman.
Rohdenburg, Günter. 1995. On the replacement of finite complement clauses by infinitives in English. *English Studies* 76(4), 367–88.
Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2), 149–82.
Rohdenburg, Günter. 2002. Processing complexity and the variable use of prepositions in English. In Hubert Cuyckens & Günter Radden (eds.), *Perspectives on prepositions*, 79–100. Tübingen: Niemeyer.
Rohdenburg, Günter. 2003. Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers. In Rohdenburg & Mondorf (eds.), 205–49.
Rohdenburg, Günter. 2005. Establishing a clause embedding hierarchy for extractions in English. Paper presented at the University of Hamburg.
Rohdenburg, Günter. 2006a. Processing complexity and competing sentential variants in present-day English. In Wilfried Kürschner & Reinhard Rapp (eds.), *Linguistik International: Festschrift für Heinrich Weber*, 51–67. Lengerich: Pabst Science Publishers.
Rohdenburg, Günter. 2006b. The role of functional constraints in the evolution of the English complementation system. In Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds.), *Syntax, style and grammatical norms: English from 1500-2000*, 143–66. Bern: Peter Lang.
Rohdenburg, Günter. 2006c. Discrepancies between the rule formulations advanced by famous linguists and their own written usage. In Markus Kötter, Oliver Traxel & Stephan Gabel (eds.), *Investigating and facilitating language learning: Papers in honour of Lienhard Legenhausen*, 47–63. Trier: Wissenschaftlicher Verlag.
Rohdenburg, Günter. 2007a. Grammatical variation in English and the formation/confirmation of linguistic hypotheses by means of Internet data. In Marianne Hundt, Carolin Biewer & Nadja Nesselhauf (eds.), *Corpus linguistics and the Web*, 191–209. Amsterdam: Rodopi.
Rohdenburg, Günter. 2007b. Functional constraints in syntactic change: The rise and fall of prepositional constructions in Early and Late Modern English. *English Studies* 88(2), 217–33.
Rohdenburg, Günter. 2013. The construction *cannot help -ing* and its rivals in Modern English. In Hilde Hasselgård, Jarle Ebeling & Signe Oksefjell Ebeling (eds.), *Corpus perspectives on patterns of lexis*, 113–32. Amsterdam: John Benjamins.
Rohdenburg, Günter. 2014. Relative clauses of reason in British and American English. *American Speech* 89(3), 288–310.
Rohdenburg, Günter. 2015. The embedded negation constraint and the choice between more or less explicit clausal structures in English. In Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds.), *Perspectives on complementation: Theory, structure and variation*, 101–27. Basingstoke: Palgrave Macmillan.
Rohdenburg, Günter & Julia Schlüter. 2000. Determinanten grammatischer Variation im Früh- und Spätneuenglischen. *Sprachwissenschaft* 25(4), 445–96.
Rohdenburg, Günter & Britta Mondorf (eds.). 2003. *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press

Ross, John Robert. 1967. Constraints on variables in syntax. PhD thesis, MIT.
Ross, John Robert. 1974. Three batons for cognitive psychology. In Walter B. Weimer & David S. Palermo (eds.), *Cognition and the symbolic processes*, 63–124. Hillsdale, NJ: Laurence Erlbaum.
Rudanko, Juhani. 1999. *Diachronic studies of English complementation patterns: Eighteenth century evidence in tracing the development of verbs and adjectives selecting complement clauses*. Lanham, MD: University Press of America.
Rudanko, Juhani. 2000. *Corpora and complementation: Tracing sentential complements in English in recent centuries*. Lanham, MD: University Press of America.
Rudanko, Juhani. 2010. Tracking and explaining variation and change in the grammar of American English: A case study with evidence from the *TIME* Corpus. In Ursula Lenker, Judith Huber & Robert Mailhammer (eds.), *English historical linguistics 2008*, vol. I: *The history of verbal and nominal constructions*, 29–44. Amsterdam: John Benjamins.
Szamosi, Michael. 1973. On the unity of subject raising. *Chicago Linguistic Society (CLS)* 9, 652–8.
Vosberg, Uwe. 2003a. The role of extractions and *horror aequi* in the evolution of *-ing* complements in Modern English. In Rohdenburg & Mondorf (eds.), 305–27.
Vosberg, Uwe. 2003b. Cognitive complexity and the establishment of *-ing* constructions with retrospective verbs in Modern English. In Marina Dossena & Charles Jones (eds.), *Insights into Late Modern English*, 197–220. Bern: Peter Lang.
Vosberg, Uwe. 2006. *Die Große Komplementverschiebung: Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.
Warner, Anthony. 1982. *Complementation in Middle English and the methodology of historical syntax*. London: Croom Helm.
Wasserman, Robert David. 1976. Theories of linguistic variation. PhD thesis, Indiana University.
Zipf, George K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S1360674316000307 Published online by Cambridge University Press