languages

MDPI

Article

Preposition Stranding vs. Pied-Piping—The Role of Cognitive Complexity in Grammatical Variation

Christine Günther

Faculty of Arts and Humanities, Universität Siegen, 57076 Siegen, Germany; guenther@anglistik.uni-siegen.de

Abstract: Grammatical variation has often been said to be determined by cognitive complexity. Whenever they have the choice between two variants, speakers will use that form that is associated with less processing effort on the hearer's side. The majority of studies putting forth this or similar analyses of grammatical variation are based on corpus data. Analyzing preposition stranding vs. pied-piping in English, this paper sets out to put the processing-based hypotheses to the test. It focuses on discontinuous prepositional phrases as opposed to their continuous counterparts in an online and an offline experiment. While pied-piping, the variant with a continuous PP, facilitates reading at the *wh*-element in restrictive relative clauses, a stranded preposition facilitates reading at the right boundary of the relative clause. Stranding is the preferred option in the same contexts. The heterogenous results underline the need for research on grammatical variation from various perspectives.

Keywords: grammatical variation; complexity; preposition stranding; discontinuous constituents

check for updates

Citation: Günther, Christine. 2021. Preposition Stranding vs. Pied-Piping—The Role of Cognitive Complexity in Grammatical Variation. *Languages* 6: 89. https://doi.org/10.3390/languages6020089

Academic Editors: Sina Bosch, Ilaria De Cesare, Ulrike Demske and Claudia Felser

Received: 29 December 2020
Accepted: 14 May 2021
Published: 18 May 2021

1. Introduction

Grammatical variation refers to phenomena where speakers have the choice between two (or more) semantically equivalent structural options. Even in English, a language with rather rigid word order, some constructions allow for variation, such as the position of a particle, the ordering of post-verbal constituents or the position of a preposition.

(1) a. She looked the number up.
b. She looked up the number.
(Lohse et al. 2004, p. 238)

(2) a. Pat talked to Chris about Sandy.
b. Pat talked about Sandy to Chris.
(Wasow and Arnold 2003, p. 119)

(3) a. [To your father] I'm even more deeply indebted.
b. [Your father] I'm even more deeply indebted to __.
(Huddleston and Pullum 2002, p. 627f)

As shown in example (1), a transitive multi-word verb can either occur as a single unit or the verb and the particle can be "split" by an intervening direct object. Example (2) illustrates the variable ordering of prepositional phrases, a further instance of variation in the verb phrase. As can be seen in example (3), preposition placement also allows for flexibility in certain types of movement contexts, where the complement of a preposition is fronted. The preposition can either be moved along with its complement, i.e., it is "pied-piped" (Ross 1967), or it can remain in situ, which means it is left stranded. Example (3a) shows the pied-piped version: The prepositional complement of the participle *indebted* is moved to the front of the clause. In (3b), only the nominal complement of the preposition is moved while the latter remains stranded in its original position following the verb.

Further grammatical variation phenomena are the placement of adpositions, the dative alternation, the position of complex adjective phrases in NP structure, and the choice between zero and overt forms. What all these phenomena have in common is that their determinants, grammatical weight probably being the most prominent one, have been linked to cognitive complexity (see, e.g., Gries 2003; Lohse et al. 2004) on particle placement, (Hawkins 2000; Wasow and Arnold 2003) on post-verbal constituent ordering, (Trotta 1998; Gries 2002; Hoffmann 2011) on preposition placement, (Berlage 2009, 2014) on adposition placement (Bresnan and Ford 2010) on the dative alternation (Hawkins 2014; Günther 2018) on attributive adjective phrases (Hawkins 2003; Rohdenburg 2003) on zero vs. overt forms. The common, underlying idea of processing-based approaches can be summarized as follows: One of the variants is associated with a higher processing cost, which may be further enhanced by a complex syntactic environment. In order to reduce cognitive complexity, speakers select an easier, or in Hawkins' (2014) terms, a more "efficient" variant. This is evident in Rohdenburg's (2003) Complexity Principle, for instance, which states that "in the case of more or less explicit constructional options the more explicit one(s) will tend to be preferred in cognitively complex environments" (Rohdenburg 2003, p. 217). Which of the variant is the more difficult or the "less explicit" one and what contributes to a cognitively complex environment depends on the individual phenomenon. In particle placement, for instance, it is often assumed that a discontinuous verb–particle combination is more difficult to process because of the close (lexical and syntactic) dependency relation between the two elements (cf. Gries 2003). In order to reduce the load, speakers use the continuous variant in syntactic environments that are associated with a higher processing load such as more complex direct objects.

Relating grammatical variation to processing strategies is certainly a promising idea but the approaches are not without problems. Many analyses put forth psycholinguistic explanations but they do not provide psycholinguistic evidence, see also (Kunter 2017) for a similar point of criticism. Instead, most of the analyses are based on corpus data, which provide only indirect evidence for processing-based hypotheses.

This paper seeks to complement the existing corpus-based studies with experimental data. Its central aim is to shed light on the question of whether the hypotheses from the corpus-based literature are supported in experimental paradigms. The paper reports two experiments, a self-paced reading task and an acceptability judgment experiment, both of which focus on preposition placement vs. pied-piping in restrictive relative clauses in English. This alternation is particularly interesting for several reasons. First, there are extensive corpus-based studies on the phenomenon, which allows a detailed comparison of experimental against corpus data. Second, psycholinguistic studies on distance dependencies have shown that, in some cases, the distance between dependent items facilitates processing (e.g., Konieczny 2000) in other cases the distance induces processing difficulties (e.g., Vasishth and Drenhaus 2011).

As will be shown, pied-piping facilitates reading at the relative pronoun and slows down reading at the right boundary of the relative clause, while a stranded preposition has the opposite effect. With syntactically simple relative clauses, the facilitatory effect and the increased processing load cancel each other out, i.e., there is no difference in reading times between the two variants for the whole stimulus. With structurally complex relative clauses, however, the cumulative reaction times show a slight advantage of the pied-piped variant. The acceptability experiment shows that stranding overall is the preferred variant.

The paper is structured as follows. Section 2 provides a discussion of previous, complexity-based approaches to preposition placement in English. The reading experiment is reported in Section 3, the rating task is reported in Section 4. Section 5 discusses the conflicting results of the two experiments. A summary and a conclusion are provided in Section 6.

2. Preposition Placement in English—Stranded vs. Pied-Piped Prepositions

According to Huddleston and Pullum (2002), preposition stranding occurs in a range of different syntactic environments, while pied-piping ("PP fronting") is more restricted. An overview is presented in Table 1.

Table 1. Preposition placement in English (adapted from Pullum and Huddleston 2002, pp. 627–28).

Construction	Stranded Variant	Pied-Piped Variant
Preposing	Your father I'm even more deeply indebted to.	To your father I'm even more deeply indebted.
Open interrogative	Who are they doing it for?	For whom are they doing it?
Exclamative	What a magnificent table the vase was standing on!	On what a magnificent table the vase was standing!
Wh-relative	He's the one who I bought it from.	He's the one from whom I bought it.
Non-*wh* relative	He's the one (that) I bought it from.	—
Comparative	Kim went to the same school as I went to.	—
Hollow clause	His performance was easy to find fault with.	—
Passive	The bed looks as if it had been slept in.	—

What should be noted is that, even though stranding is licit in more syntactic constructions, this variant is considered "grammatically incorrect" (Pullum and Huddleston 2002, p. 627) from a prescriptive perspective, which stands in sharp contrast to the extensive use of stranding in everyday Standard English.

Several approaches to preposition placement as a grammatical variation phenomenon have established a link to cognitive complexity such as (Trotta 1998; Gries 2002; Hoffmann 2011). Based on corpus data, these authors relate (some or all of) the determinants to processing. Interestingly, there is no general agreement on which variant, the stranded or the pied-piped preposition is the more difficult one, as will become clear below.

Gries (2002) assumes that (primarily) production costs determine the choice between the stranded and the pied-piped variant. The stranded variant is seen as the more difficult one because of a filler-gap dependency. "Filler" refers to a displaced element, such as a *wh*-phrase, and "gap" refers to the position the element assumes in a canonical structure or its base position. Filler-gap structures are associated with processing cost, as illustrated in the following quote from (Hawkins 1999):

Filler-gap dependencies are difficult structures to process, and they are characterized by a heightened processing load and a constant effort to relate the filler to its appropriate gap. Identifying the gap is not easy. It is an empty element with no surface manifestation and its presence must be inferred from its immediate environment. At the same time, the filler must be held in working memory, all the other material on the path from filler to gap must be processed simultaneously, and the gap must be correctly identified and filled. (Hawkins 1999, p. 246f).

Gries' analysis includes multiple potential determinants, ranging from morpho-syntactic to discourse-functional variables, all of which are linked to processing complexity. He predicts that stranding will be infrequent in contexts that are difficult to process already, because as the more difficult variant, it would add to an already high level of processing cost. A multifactorial regression model identifies properties of the bridging structure, i.e., the structure between the displaced element and the extraction site, transitivity of the verb, the voice of the verb as well as modality as predictors. Longer or syntactically more complex bridging structures, transitive verbs, passive constructions and written language show a preference for pied-piped prepositions. Modality is argued to be an effect of prescriptive rules, but the other factors are considered to confirm the processing-based line of reasoning, because complex bridging structures require more processing effort and a direct object of the transitive verb increases the distance between the preposition and its complement. Passives are also considered to impose processing load on the speaker, which is why speakers avoid stranding in these contexts. However, Gries acknowledges that this connection is not straightforward and leaves room for further examination.

This brief summary already touches upon some of the problematic aspects in corpus-driven investigations of preposition placement. First of all, there is a lack of independent, psycholinguistic evidence for what contributes to processing difficulties. Second, the claim that stranding is the more demanding variant is not convincingly motivated: Gries relates it to Hawkins' explanations on how filler-gap dependencies have an effect on grammatical variation. However, the pied-piped variant also contains a gap (due to the displaced *wh*-element). It is true that Hawkins (1999) argues for the stranded version to be more difficult in the following example:

(4) Which student did you ask Mary about?
(Hawkins 1999, p. 247)

However, this is an effect of ambiguity in this particular structure. As shown in example (4′), there are two possible gaps for *which student*.

(4′) Which student$_i$ [did you ask (O$_i$) Mary about O$_i$]?

Hearers first interpret *which student* as the object of the verb *ask* and then have to revise this analysis once they encounter *Mary*, which results in a processing difficulty. Importantly, this effect is not a general property of pied-piping vs. stranding, it rather depends on the verb.[1] With prepositional verbs that are unambiguous with regards to transitivity, such as *rely on*, the issue does not arise.

(5) Which student$_i$ [did you rely on O$_i$]?

What is more, the pied-piped version can also contain a misleading gap site, if the verb allows for understood objects as is the case with *read*.

(6) About which topic$_i$ [did you read (O$_i$) a book O$_i$]?

As (6) shows, *read* could be interpreted as a prepositional verb that does not have an overt direct object and could be mistaken as the clause-final element.

This discussion shows that it is not entirely convincing to generally consider preposition stranding as the more difficult variant. A filler-gap dependency holds in both variants of preposition placement. It could be argued, though, that stranding adds a discontinuity—the preposition is separated from its complement, resulting in a distance dependency between two elements, which increases with a longer bridging structure. This could be taken as an explanation of why longer bridging structures tend to occur with the pied-piped variant, as Gries (2002) finds (cf. also (Trotta 1998, p. 207)). Yet, the pied-piped variant also includes a distance dependency because a preposition is separated from its licensing verb. An anonymous reviewer pointed out that the two different gap sites do not have to be associated with the same cognitive load. This is true, especially since there are different dependency relations that can hold between heads and their complements, such as the lexical and syntactic relation between a prepositional verb and its complement. As will become clear in the discussion of Hoffmann's (2011) analysis below, this aspect is also highly relevant for the distributional properties of the construction.

In Trotta's (1998) data from the Brown corpus, pied-piping is more frequent than stranding in relative clauses, but stranding is more frequent than pied-piping in exclamatives and interrogatives. The latter is related to a discourse function of the *wh*-element, which provides a clear indication of the illocutionary force. The preference for pied-piping in relative clauses, though, as well as the effect of distance between filler and gap is related to processing: If the distance becomes greater, "the discontinuous PP becomes difficult to process as a single constituent" (Trotta 1998, p. 207). However, Trotta himself considers this reason "speculative".

Hoffmann (2011) presents a study of the International Corpus of English (ICE), which is complemented by two acceptability judgment tasks. The corpus analysis considers a range of different factors such as the nature of the PP, the syntactic context and genre.

In Hoffmann's data, free clauses and interrogatives (main and embedded) show a higher proportion of stranded prepositions, while relative clauses and clefts show a higher number of pied-piped prepositions. Overall, pied-piping is more frequent than preposition stranding. The type of displaced element also plays a role with some elements displaying a

1 The interpretation might also depend on the subject. If the object is changed to an inanimate entity that is not in the subcategorization frame of *ask*, the first gap is rather unlikely. (i) Which shirt$_i$ [did you ask $^{\#}$(O$_i$) Mary about O$_i$]?

higher amount of stranded and others a higher amount of pied-piped prepositions—*what*, for instance, is rarely found with pied-piping while there are only few stranded tokens for *which*. The phenomenon is most prominent in the verb phrase, where pied-piping is more frequent. As expected, pied-piping is the preferred variant in more formal contexts (such as printed/edited texts). Stranding is much more frequent than pied-piping in the more informal text types (such as private dialogue). Close relations between prepositions and their licensing verbs show a preference for stranding. With prepositional adjuncts, pied-piping is more frequent.

Relative clauses are the context in which preposition placement is most frequent in Hoffmann's data (2011, p. 158). In this clause type, pied-piped prepositions are more frequent than stranded ones. In a separate analysis of the relative clause data, Hoffmann factors in the syntactic complexity of the relative clause, based on Lu's (2002) complexity metric and restrictiveness. For non-restrictive clauses, the proportion of stranded prepositions is higher than for restrictive ones. According to Hoffmann (2011, p. 170) "non-restrictive relative clauses are not necessary for the identification of the reference of the antecedent NP. Therefore, the filler-gap identification process in non-restrictive relative clauses is less complex [...], which also accounts for the favouring stranding effect".

The relative clause data reveal an interaction of complexity and PP type. Interestingly, complex clauses show a higher proportion of preposition stranding if there is a close lexical dependency between verb and preposition, as is the case with prepositional verbs. In this particular context, stranding is even slightly more frequent than pied-piping.

Hoffmann conducts a series of acceptability judgment experiments to zoom in on some of the above findings. In a magnitude estimation task on relative clauses in British English, he also investigates the effect of bridging structures. He uses prepositional verbs such as *apologize for*, *ask for*, *belong to*, *call on*, etc., and bridging structures such as *you claimed*, *I imagined*, *I read* and *you pointed out*. Interestingly, stranding, overall, is preferred over pied-piping, except for the relativizer *whom* in simple clauses. For complex clauses, stranding receives higher scores for all relativizers (*that*, *who*, *whom*, *zero*). This is again related to the lexical dependency relation: According to Hoffmann, stranding "facilitates the integration of lexicalized verb-preposition structures" (Hoffmann 2011, p. 202), i.e., it also comes with a processing advantage.

Radford et al. (2012) also investigate preposition placement in English relative clauses from an experimental perspective. Conducting two acceptability judgment tasks, a speeded and an untimed one, they examine speakers' reactions to four variants of preposition placement—preposition stranding, pied-piping, preposition copying and preposition pruning. In preposition copying, a preposition is realized at the displaced element as well as the gap, i.e., it is doubled. In pruning, in contrast, no preposition is realized. The central aim of the study is to identify the source of copying and pruning. However, even though the focus is on a different aspect, the study provides insights on preposition stranding and pied-piping that are relevant for the present paper and will hence be summarized here. The first experiment is an acceptability judgment experiment examining the general acceptability of the four variants in restrictive relative clauses. Interestingly, in this experiment, preposition stranding and preposition pruning both elicit lower ratings than pied-piping and preposition copying. According to the authors, this could reflect a potential influence of prescriptive rules or "that pied-piping generally facilitates the processing of wh-clauses—or perhaps both" (Radford et al. 2012, p. 413). The same materials are tested in a second experiment, which is a speeded forced-choice task. Here, preposition stranding, copying, and pied-piping result in a comparable proportion of positive responses, the response latencies are similar as well. For the pruning condition, in contrast, a high proportion of responses is negative and the reaction times are slower. The divergent results for stranding support the hypothesis of a prescriptive influence on variation, suggesting that prescriptive rules have a smaller effect when processing pressure is higher, as is the case in the speeded task. The second experiment also reveals that prepositions facilitate processing in *wh*-relative clauses: The absence of a preposition results in significantly fewer positive responses as well as slower reaction times. However, no significant difference was found between stranded and pied-piped prepositions. What is more, potential facilitation effects of stranding and of pied-piping, as discussed above, do not add up, because the copying condition neither elicited a higher proportion of positive responses nor faster reactions than pied-piping and stranding in Radford et al.'s second experiment.

To sum up: Gries (2002) considers stranded prepositions the more cognitively more demanding variant, irrespective of the clause type. Trotta (1998) assumes that preposition stranding induces processing difficulties in relative clauses as well as in contexts of a greater distance between the preposition and its complement. Hoffmann (2011) also finds that pied-piping is the more frequent variant in relative clauses except for complex relative clauses with prepositional verbs. This frequency effect is mirrored in a subsequent acceptability judgment test: Speakers prefer preposition stranding over pied-piping in relative clauses with prepositional verbs. Hoffmann considers stranding the less demanding variant in this particular context. In Radford et al.'s (2012) study, however, stranding in relative clauses does display processing advantages over pied-piping.

The discussion above demonstrates that the phenomenon of preposition placement is subject to many different factors and that the question of which of the variants is more difficult to process and why is not entirely clear. The subsequent sections report two experiments on preposition placement in *wh*-relative clauses in English. The central question is whether one of the variants is easier to process, and if so, whether it is preposition stranding or pied-piping. To this end, a self-paced reading experiment was conducted. A follow-up question is at which point in the structure a processing load is induced—as illustrated above, there are two relevant dependency relations: one between the verb and the preposition and one between the preposition and its complement. This implies that in pied-piping contexts, the verb is separated from its prepositional complement whereas in stranding contexts, the preposition is separated from its complement. At each of these points, a difficulty could arise. A further research question addresses the role of structural complexity, which is operationalized as a bridging structure in the relative clause. The question is whether a more complex syntactic environment, which also results in an increase in distance between the filler and the gap, adds to a potential processing difficulty.

A second experiment uses the same set of materials in a split rating task, an acceptability judgment paradigm. Hoffmann's data show that the distribution in corpus data of the two variants only partly aligns with speakers' judgments, which makes it interesting to see if the option that is easier to read is also the preferred variant.

3. Experiment 1—Self-Paced Reading Experiment

The idea behind a self-paced reading paradigm (SPRT) is that reading times are a window to cognition and that processing difficulties are reflected in reading latencies (see, e.g., Just and Carpenter 1980).

As pointed out before, the focus of the self-paced reading experiment is on preposition stranding versus pied-piping and the effect of bridging structures in *wh*-relative relative clauses with prepositional verbs, i.e., verbs that license a particular preposition. The relative clauses are finite and restrictive. They contain a prepositional verb such as *agree with* and *search for*.

3.1. Factors, Materials and Design

The first factor is ORDER, which operationalizes the different possible positions of the preposition. It has two levels (stranded/pied-piped). This factor is crossed with the factor COMPLEXITY, which also has two levels (simple/complex). A complex condition was operationalized as the presence of a bridging structure. Crossing the factors ORDER and COMPLEXITY results in four different conditions illustrated in Table 2. For a list of all critical experimental items, see Appendix A.

Table 2. Example of experimental stimuli in self-paced reading task.

Condition	Example
Stranding, simple clause	Sam provides the answer which Ashley searched for really desperately.
Pied-piping, simple clause	Sam provides the answer for which Ashley searched really desperately.
Stranding, complex clause	Sam provides the answer which I say Ashley searched for really desperately.
Pied-piping, complex clause	Sam provides the answer for which I say Ashley searched really desperately.

The relative clause in the experimental stimuli in contains a clause-final adjunct, which was added for methodological reasons. The construction under investigation is a word order phenomenon that ends in the critical element—a preposition plus a gap in the stranded variant, example (7a), or just a gap in the pied-piped variant, example (7b), which makes the verb the final (overt) element in this case.

(7) a. Sam provides the answer [which] Ashley searched for __.
b. Sam provides the answer [for which] Ashley searched __.

Thus, the relative clause has different elements at the right boundary across the different ordering conditions. This implies that the reaction times at this point cannot be compared across conditions—a preposition is likely to be read faster than the main verb of a clause, irrespective of preposition placement. To provide a point of measurement that is uniform across all conditions, the adverbial phrase was added. It contains two adverbs to provide two points of measurement that are constant across the different conditions.

The experiment contained 24 experimental stimuli, 6 for each condition, 32 stimuli for an experiment on particle placement, as well as 46 fillers. The 46 fillers were followed by a *yes–no* comprehension question, which was used to ensure that participants stayed focused. The questions followed the fillers in order to draw the participants' attention away from the critical items.

All sentences were counter-balanced across four lists, resulting in a mixed design. This implies that there were four groups of participants. Each list contained each sentence in one condition, i.e., 6 different sentences of each condition. Thus, each participant read every sentence only once and was exposed to all four conditions. For each participant, the order of the elements in the list was presented in a different pseudo-random order.

The items were presented word by word in a moving window manner.

3.2. Predictions

A higher processing load is reflected in longer reading times. As to the effect of ORDER, there are different predictions. Based on Gries' (2002) analysis, stranding is expected to be more difficult to process. Based on Hoffmann's (2011) data, pied-piping is expected to be the more difficult variant. The findings from Radford et al., in contrast, suggest that neither variant is more difficult.

If there is a difficulty, it should result in slower reactions for the words that follow the critical region (see Table 3 in Section 3.4). In addition to local difficulties, the reaction for the whole stimulus should require more time for the variant that is associated with a higher processing load.

For the effects of COMPLEXITY, there are different predictions as well: If stranding is more difficult because of the distance dependency between the preposition and its complement, the stranded variant should be more difficult to read with complex relative clauses than with simple ones. If pied-piping is more difficult because of the distance dependency of the verb and the preposition it licenses, the pied-piped variant should be more difficult to read with complex relative clauses than with simple ones. Hence, the factors COMPLEXITY and ORDER are expected to interact.

3.3. Participants and Procedure

The experiment was conducted at the University of Edinburgh. A total of 51 students (native speakers of British English) participated in the study. They were paid for their participation. The age range is 18–39, the mean age is 20.65 years. There were 10 male and 41 female participants.

Participants received oral and written instructions. They were given 8 practice items to familiarize themselves with the procedure before they started the experiment. The experiment was run using E-Prime 2.0. The average duration was 13 min, ranging from 10:03 for the fastest to 16:41 for the slowest participant.

3.4. Data and Analysis

One participant was excluded from the final dataset because he or she had misunderstood the task and read the stimuli out. Data from five participants were excluded because these participants showed an accuracy of less than 80% in the comprehension questions. The final dataset contains data points from 45 participants, which results in a total of 1080 observations.

Table 3 depicts the different points of measurement.

Table 3. Points of measurement in Experiment 1, right sentence boundary.

Preposition			PRECRITICAL	CRITICAL	INTENSIFIER	ADVERB
Stranded	which	Ashley	searched	for	really	desperately.
Pied-piped	for	which	Ashley	searched	really	desperately.

The reaction time of the element following the critical region, INTENSIFIER, is taken as the first point of measurement, because it is uniform across all conditions and might show spill-over effects of processing difficulty in the critical region. The next point of measurement is ADVERB. It is analyzed to look for potential wrap-up effects, processing difficulties that display at the end of a clause (cf. Mitchell and Green 1978). The discussion of the literature in Section 2 indicated that stranding and pied-piping might result in difficulties at very different points throughout the structure. If this is the case, an answer to the question as to whether one of the variants is associated with a higher cognitive load is not possible on the basis of two individual measuring points following the critical region. If speakers choose a construction because it is easier to process, as claimed in the corpus-based literature, the advantage should be evident in the structure as a whole when compared to its counterpart. Thus, in order to avoid conclusions that are based on local difficulties (or advantages) only, SENTENCE was introduced as a further dependent variable. It contains the cumulative reaction times of the entire stimulus and serves to detect global difficulties.

The analyses in Sections 3.6.1–3.6.3 will demonstrate that it is the case that the stranded variant facilitates processing at the intensifier following the relative clause but shows slower reaction times for the whole sentence in the complex condition. Thus, obviously, the stranded variant induces difficulty earlier on in the stimulus. The factor ORDER creates structures that differ at two points: the left boundary of the relative clause (noun plus relative pronoun vs. noun plus preposition) and the right boundary (verb plus preposition vs. verb plus intensifier). As the facilitatory effect of stranding at the right boundary is overridden, it is likely that pied-piping facilitates reading at the left boundary. In order to shed more light on this, the reaction time at the relative pronoun *which*, REL, was included as a fourth dependent variable. This point differs from the previous because it occurs in two different positions in the stimulus (it is word 5 in the stranded variant and word 6 in the pied-piped construction), but it is the first word in the relative clause that occurs in both variants (as argued in Section 3.1, comparing the reaction times of two different lexical items, a preposition and a pronoun in this case, would be problematic).

Since all RT-variables showed a right-tailed distribution they were log-transformed to the base of 2. For each dependent variable, outliers were removed before model fitting. To this end, observations that exceeded 2.5 standard deviations from the mean were excluded both by ITEM and by SUBJECT (Baayen and Milin 2010). This resulted in a loss of 1.57–3.61% of the data points.

3.5. Statistical Models and Effects

For the analysis, linear mixed-effects models (Baayen et al. 2008) were chosen, using the *lme4* (Bates et al. 2015) and *lmerTests* package (Kuznetsova et al. (2017)) for R. These regression models allow for variable slopes and intercepts, which makes them highly suitable for data that contain repeated-measure variables such as the experimental stimuli or the participants, who contributed several data points.

The following effects were included:

- ORDER: The preposition can be *pied-piped* or *stranded*
- COMPLEXITY: The relative clause is either *simple* or *complex*
- TRIAL: The number of the trial, or in other words, the position of a sentence within the experiment
- ITEM: The different experimental stimuli
- LOGCRITICAL: The log-transformed RTs at the critical region
- LOGPRECRITICAL: The log-transformed RTs at the element preceding the critical region
- SURPRISALGAP: The relative degree of surprisal of the intensifier
- SURPRISALREL: The relative degree of surprisal of the relative pronoun
- SUBJECT: The participant

While most of the variables are self-explanatory, some need further discussion. Several were added to the list in order to control for influences on the dependent variables apart from the four different conditions. The RT of one or two words preceding the point of measurement was included (see Bartek et al. 2001), to control for spill-over effects and differing material in the critical region across the different experimental stimuli. A further possible influence on the reaction time is to what extent an item is expected at a certain point—items that are more predictable will be easier to process than those that are less predictable (e.g., Levy 2008). For this reason, surprisal was introduced as a further variable. It is operationalized as the negative binary log of the probability of a word after a previous word (e.g., Rühlemann and Gries 2020). There are two surprisal values: SURPRISALREL, which is the negative binary log of the bigram frequency of *which* and the word preceding it in the British National Corpus divided by the BNC-frequency of the first word, i.e., the noun or the preposition. SURPRISALGAP is computed as of the negative binary log of the BNC-frequency of the intensifier and the word preceding it plus 1[2], divided by the BNC-frequency of the first word.

Since the effects of ORDER and COMPLEXITY might change within the course of the experiment—participants might get used to a particular construction—an interaction term was included for TRIAL, ORDER and COMPLEXITY.

The first model, M0, contained all the fixed effects that were considered relevant for a particular point of measurement as well as SUBJECT and ITEM as random intercepts. While participants are likely to have a different reading pace in general, their reaction to the different conditions might vary more or less drastically. The reading pace might also slow down or speed up during the experiment, which is why the null model had random slopes for COMPLEXITY, ORDER and TRIAL for SUBJECT. Similarly, the reading times for stimuli might show different slopes due to a lexical bias, which is why M0 also contained random slopes for COMPLEXITY and ORDER for ITEM. The first steps in the model selection process included the stepwise reduction of an overly complex random effect structure, by removing those components that accounted for less than 5% of the variance in principal component

[2] Some bigrams were not attested, which would have resulted in undefined values for log(0).

analysis. As a second step, the models were subjected to model criticism (cf. Baayen and Milin 2010) because the residuals showed non-normality. Finally, the non-significant fixed effects were removed step by step, starting with the effect that had the highest *p*-value. The models for the different steps in the reduction process as well as the final models are summarized in Appendix B.

3.6. Results

Table 4 provides an overview of the raw reaction times across the different points of measurement.

Table 4. Mean reaction times and standard deviations in milliseconds for four conditions across points of measurement in Experiment 1.

	Points of Measurement							
	REL		INTENSIFIER		ADVERB		SENTENCE	
Condition	**Mean**	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Stranding, simple clause	391	157	407	170	506	290	4104	1456
Pied-piping, simple clause	375	150	416	163	494	259	4070	1459
Stranding, complex clause	404	178	427	171	528	278	4999	1780
Pied-piping, complex clause	382	155	435	191	533	284	4865	1756

3.6.1. Spill-Over Region

In the model for the reaction times in the spill-over region, INTENSIFIER, two of the preceding reaction times (LOGCRITICAL and LOGPRECRITICAL) were included as controls because of the very different categories across the two ordering conditions (see Table 3). Not surprisingly, these two variables show a rather strong correlation (rs = 0.7, p = 0), which might result in collinearity issues. A principal components analysis suggested that each of them accounted for a considerable proportion of the variance, which is why the principal components, PREVPC1 and PREVPC2, were integrated as predictors instead of LOGCRITICAL and LOGPRECRITICAL. TRIAL, ORDER and COMPLEXITY were added as interaction terms in M0. The null model also included SURPRISALGAP as predictors. The random effect structure of the null model was as laid out above. The final model (Table A7, Appendix B) had a random intercept for SUBJECT, as well as random slopes for ORDER and COMPLEXITY.

ORDER has a small but significant effect at this point of measurement (t = 2.894, p = 0.006), but COMPLEXITY does not. There is no interaction effect, either. Interestingly, the reactions are slower for the pied-piped variant. This means that the stranded variant results in facilitation in the spill-over region at the *ly*-adverb. Figure 1 illustrates the effect of ORDER. Please note that the plot shows the binary log of the dependent variable. A back-transformation of the predicted values (pied-piping: 8.619808, stranding: 8.581056) reveals a difference of around 10.5 ms.

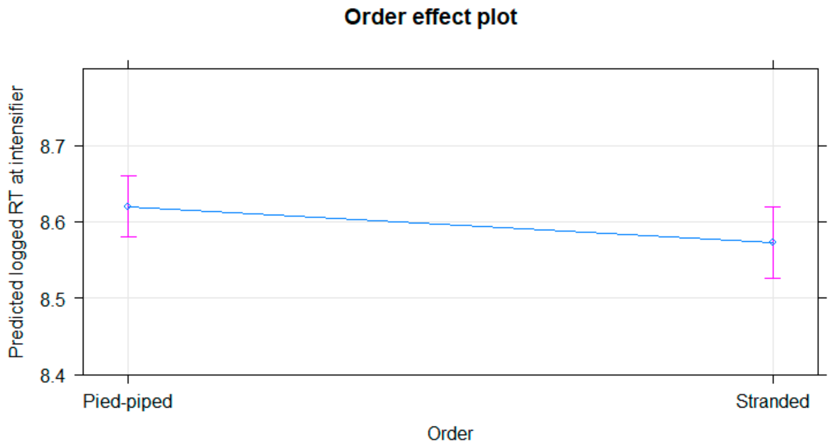

Figure 1. Effect of ORDER in the spill-over region of Experiment 1.

3.6.2. Wrap-Up Region

The next dependent variable is the reading time at the sentence-final adverb, the wrap-up region. Since the point of measurement is preceded by the same adverb across the different conditions, only one previous RT was included. The other effects in the null model were identical to the null model described above. The final model contains a random intercept for SUBJECT and for ITEM. The fixed effects are summarized in what follows (see Table A14, Appendix B for details).

This time, COMPLEXITY has a small but significant effect (t = 3.108, $p = 0.002$): The model predicts an advantage for the simple conditions. The difference of the back-transformed predicted values (complex: 8.845969, simple: 8.783326) is about 20.5 milliseconds. ORDER does not have an effect. The effect is illustrated in Figure 2.

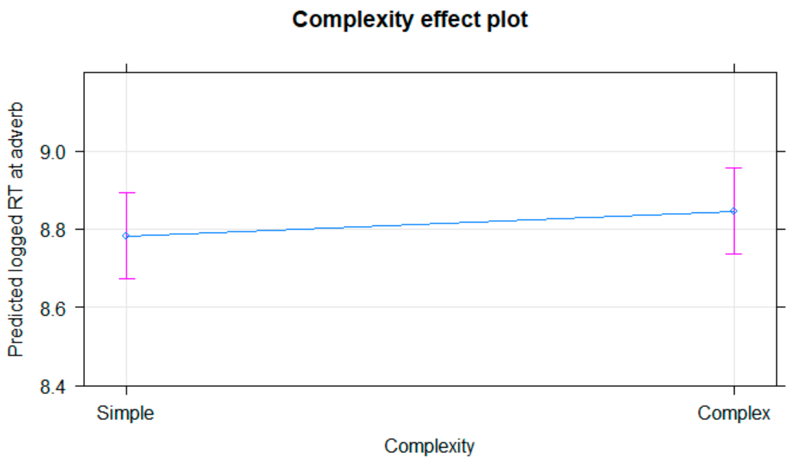

Figure 2. Effect of COMPLEXITY in the wrap-up region of Experiment 1.

3.6.3. Relative Pronoun

For the reaction time at the relative pronoun, the principal components of the RTs of two elements preceding the pronoun were used instead of said RTs, because, again, there was a strong correlation between them (rs = 0.75, $p = 0$). This time SURPRISALREL was included, while COMPLEXITY was not because this factor describes a difference that occurs *after* this point of measurement. The final model has a random intercept for SUBJECT with a random slope for ORDER. It is summarized in Table A16, Appendix B.

ORDER has a significant effect (t = 2.376, $p = 0.023$). At this point, there is a very slight advantage of the pied-piped variant (8.457463 vs. 8.498022 on the binary log scale), which is about 10 milliseconds faster. The effect is illustrated in Figure 3.

Figure 3. Effect of ORDER at the relative pronoun.

3.6.4. Cumulative Reaction Times

For the cumulative reaction times, two separate models were fitted for the data in the simple and the data in the complex condition, because the two additional words of the bridging structure necessarily result in longer reaction times for the complex condition. The null models contain SURPRISALGAP und SURPRISALREL as fixed effects in addition to ORDER.

In the final model for the simple condition, there is no effect of ORDER, which is why it is not reported here (it is included as Table A20 in Appendix B). The final model for the complex condition (Table A25, Appendix B) has a random intercept for SUBJECT and for ITEM. ORDER has an effect (t = 2.755, $p = 0.006$), which is shown in Figure 4. The model predicts a logged RT of 12.15476 in the pied-piped variant and 12.19383 in the stranded, which means that the stranded variant is predicted to take about 125 milliseconds longer to read.

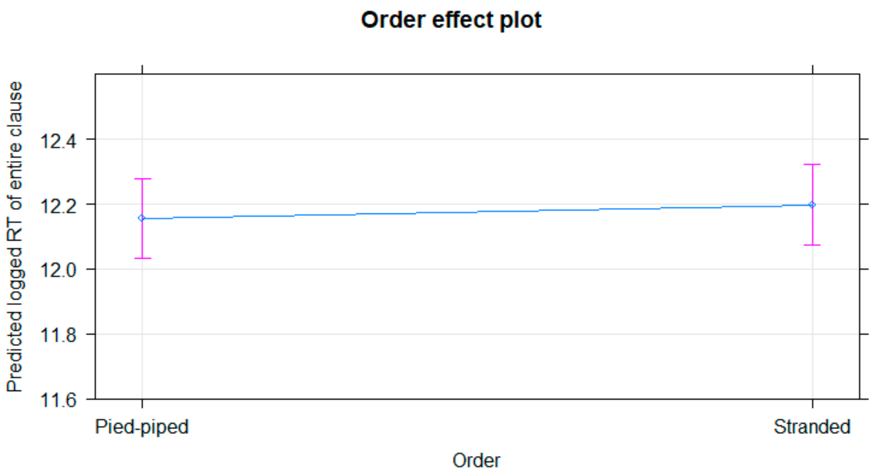

Figure 4. Effect of ORDER in cumulative reaction times of complex clauses in Experiment 1.

3.7. Discussion

To sum up: Preposition stranding results in faster reading times at the right boundary of the relative clause, whereas pied-piping results in faster reading times at the left. In the wrap-up region, there is no effect of ORDER. For the complex condition, there are faster reaction times for pied-piping in the cumulative perspective. COMPLEXITY has an effect in the wrap-up region, where complex relative clauses result in slower reading times.

The analysis yielded heterogeneous results, which means there is no straightforward answer as to whether the hypotheses can be maintained or need rejection, especially since the effects are rather small. First of all, COMPLEXITY does not interact with ORDER. Not surprisingly, the additional bridging structure increases the processing load, which is evident in the wrap-up region, but this does not seem to contribute to the processing cost associated with one of the two preposition placement variants. However, an indirect effect can be witnessed in the cumulative RTs, because ORDER here only has an effect in the complex clauses.

The example stimulus is repeated here to illustrate the effects of ORDER. (8a) shows the stranded variant, (8b) shows the pied-piped preposition. The underlined word is the point of measurement that is read faster in that ordering condition compared to the other, the double underlining indicates slower reading in that condition.

(8) a. ... the answer which Ashley searched for really desperately.
b. ... the answer for which Ashley searched really desperately.

The effects of ORDER show opposite directions. While stranding facilitates processing at the intensifier, the spill-over region (*really* in example (8a)), pied-piping facilitates processing at another, the relative pronoun *which* in (8b). The facilitation effect of stranding in the spill-over region, i.e., at the intensifier following the preposition, can be explained with the close relationship between the verb and the preposition. Once readers have encountered the *wh*-complement of the preposition as well as the verb, they will expect to see a particular preposition next. From an expectation-based processing perspective (e.g., Levy 2008), the facilitation effect here thus does not come as a surprise. Similarly, a pied-piped preposition facilitates reading at the relative pronoun. These two effects cancel each other out in the simple condition, i.e., stimuli without a bridging structure, so that the cumulative reading times show no significant effect of ORDER. In the complex condition, in contrast, the facilitatory effect of pied-piping persists once the reaction times of the whole stimulus are analyzed. It amounts to an overall advantage of 125 milliseconds in this variant.

The explanation proposed here gives rise to a question, though: If predictions of upcoming material are said to speed up reading, why did the two surprisal variables not have a significant effect? This is likely an effect of how surprisal was operationalized here. Being based on bigram frequencies, the variable only considers one previous word. Expectations, however, can also be activated by various words (or constituents) that precede a certain element, because they narrow down the range of options. What is more, SURPRISAL is calculated for lexical items and not for the syntactic category, i.e., this encodes the probability to encounter a string of two particular words such as *answer which* and not the probability to see, for instance, any relative pronoun after any noun.

4. Experiment 2—Split Rating Task

A split rating task (Bresnan and Ford 2010) is a judgement task that contrasts two alternatives that have to be rated according to naturalness. Participants are asked to compare the two alternatives and distribute 100 points between them to express their rating, e.g., 50/50 if there is no difference, 10/90 if the second variant is much more natural than the first, etc. Every combination that adds up to one hundred is possible which means the only option not available is to reject both variants, as Bresnan and Ford (2010, p. 186) point out.

4.1. Factors, Materials and Design

As in the previously reported experiment, preposition placement in *wh*-relative clauses with prepositional verbs was investigated. The factors COMPLEXITY and ORDER were crossed, each having two levels.

This experiment is based on the same material as the self-paced reading experiment, i.e., it contained the same critical stimuli and fillers in order to establish one-to-one comparability. The comprehension questions were left out. The sentences were presented as minimal pairs that contrasted the stranded and the pied-piped variant, i.e., two simple relative clauses or two clauses including a bridging structure were displayed at the same time. The first sentence was the relevant one for the analysis, the second just served as a reference point. This implies that participants did not rate 24 critical sentences but 24 *pairs* of critical sentences. The experiment also adopted the random orders for each participant from the previous experiment, i.e., the first sentence of the pair occurred in the same trial as it did in the SPRT.

4.2. Predictions

As illustrated in Section 2, there is no general consensus on which of the two variants in preposition placement can be considered the more demanding one and to what extent structural complexity enhances cognitive complexity. Gries (2002) corpus study shows that for an increasing distance between the displaced element and the verb, there is a higher proportion of pied-piped prepositions. Hoffmann (2011) corpus study, in contrast, finds that for prepositional verbs in relative clauses, an increase in structural complexity correlates with a higher proportion of stranded prepositions, whereas for simple clauses the proportion of pied-piping is higher. This distribution is only partly reflected in the ratings from Hoffmann's magnitude estimation experiment on preposition placement in relative clauses with prepositional verbs. A complex clause increases the preference for stranding, which is comparable to the corpus findings, but unlike in the corpus data where pied-piping was the more frequent variant for simple relative clauses, stranding is also preferred in the simple condition in the rating experiment (except for the relative pronoun *who*). Radford et al. (2012) judgment data show divergent results: In an untimed experiment, there are higher ratings for pied-piping than for stranding, but in a speeded experiment the two variants are on a par. The reading experiment reported in Section 3 above has shown that pied-piping is the variant that results in slightly fasting reading times overall in the context of complex relative clauses.

In the light of the above, the predictions for Experiment 2 are not entirely straightforward. On the one hand, the experiment is based on the same materials as Experiment 1, which identified stranding in complex clauses as the more demanding variant. If speakers prefer constructions that are associated with a lower cognitive load, pied-piping should receive higher ratings than stranding in complex relative clauses. On the other hand, Experiment 2 is a rating experiment contrasting two sentences, and hence quite similar to Hoffmann's magnitude estimation experiment, even though the reference sentence here is a variant of the critical one and it the experiment does not contain ungrammatical fillers. As in Hoffmann's experiment, stranding was the preferred option and even more so in complex relative clauses, the ratings here can be expected to reflect a similar pattern. Since the task is to rate constructions according to how natural they sound, prescriptive rules are likely to play a minor role.

4.3. Participants and Procedure

There were the same participants as in the previous experiment. Experiment 2 took place on the same day as Experiment 1. Again, E-Prime was used as software. As in the reading experiment, there were oral and written instructions as well as a set of four practice items before the actual experiment started. In order to avoid the problem of potential miscalculations, only the value for the first sentence had to be entered and confirmed. The difference to 100 occurred automatically.

4.4. Data and Statistical Models

As in the self-paced reading experiment, linear mixed-effects models were fitted. This time, the full dataset was used, i.e., no participants were excluded, but the set was again scanned for outliers that were further than 2.5 standard deviations away from the mean. This resulted in 19 out of 1216 observations being removed, i.e., a loss of 1.56% of the data points. The effect structure of the null model was much simpler than the one for the SPRT data, since reaction times and surprisal variables were excluded, leaving a three-way interaction of TRIAL, ORDER and COMPLEXITY. The procedure of model selection was the same as in Section 3.5. Summaries of the fixed effects in the different models throughout the selection process are provided in Appendix B, Tables A26–A29.

4.5. Results

The final model had a random intercept for SUBJECT and ITEM, as well as random slopes for ORDER with both factors. It is summarized in Table A29 in Appendix B.

ORDER has an effect, but there is no interaction with COMPLEXITY. The stranded variant receives significantly higher ratings than the pied-piped variant (t = 4.483, $p = 0$). The effect is illustrated in Figure 5.

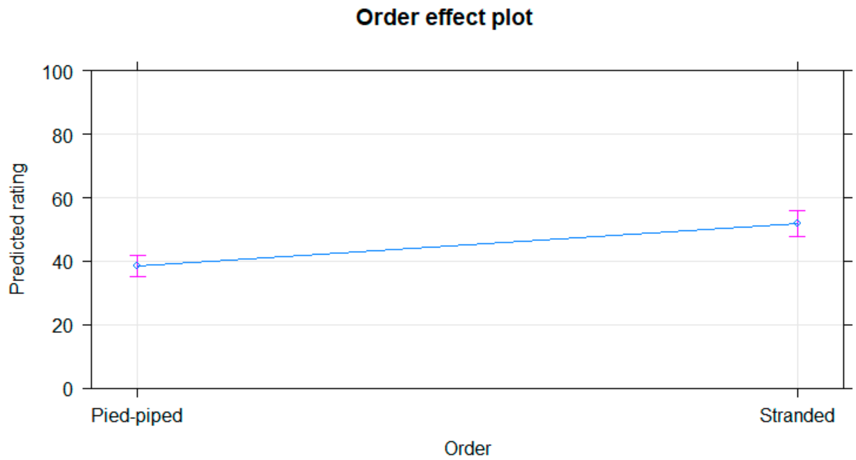

Figure 5. Effect of ORDER in Experiment 2.

COMPLEXITY displays an interaction with TRIAL. As can be seen in Figure 6, in the earlier trials items with complex clauses receive lower ratings than those with simple ones. The slope for complex clauses rises while the slope for simple ones falls with an increase in trial number, so that the advantage of simple clauses is cancelled out over time.

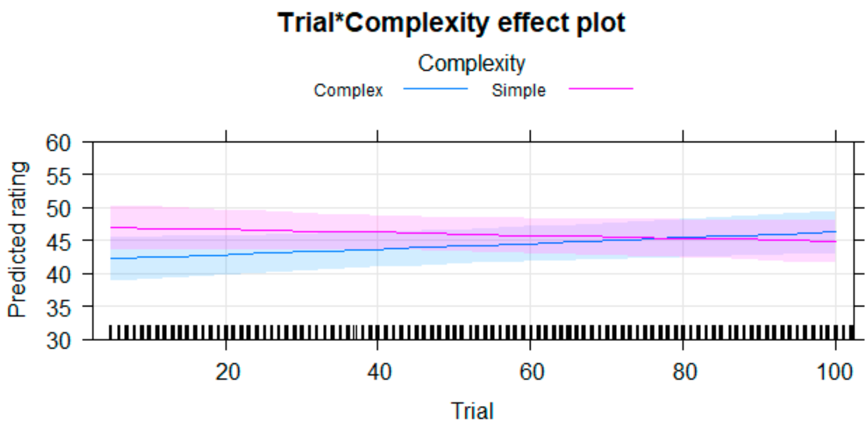

Figure 6. Interaction of TRIAL and COMPLEXITY in Experiment 2.

4.6. Discussion

Overall, stranding is the preferred variant. Even though ORDER does not interact with COMPLEXITY, the results are very similar to those in Hoffmann's (2011) magnitude estimation experiment. However, they stand in contrast to the results in Experiment 1, where the variant that receives the highest ratings here, stranding, displays the longest reading times. As the two experiments reported here used the same materials, this divergence cannot be attributed to differences in stimuli and hence has to follow from the type of task. This matter is discussed in detail in the general discussion in Section 5.

As an anonymous reviewer pointed out, priming could be an issue in the rating task because the participants saw the same experimental stimuli as in the previously conducted reading experiment. Of course, this is not entirely unproblematic but this procedure allows us to compare the data from the two different experiments in a very direct way. Since all participants saw all of the items twice and not only a subset of them, priming should affect all items. Still, the results are in line with the results of previous acceptability judgment tasks on this phenomenon and hence it is unlikely that there is a huge influence by the first experiment on the second.

5. General Discussion

The most interesting aspects about variation between stranded and pied-piped prepositions in English are the differences across the studies and the partly conflicting results. The idea that an increased processing load associated with one variant will be mirrored in lower frequencies in corpora, slower reactions in processing studies as well as lower ratings in acceptability judgment experiments is much too simple. The two experiments reported here and the findings from corpus studies indicate that, at least for preposition placement in relative clauses, it neither seems to be the case that speakers (or writers) necessarily avoid constructions that require more reading time, nor that readers prefer constructions which are faster to read.

Hoffmann (2011) corpus analysis found that in British English relative clauses, a pied-piped preposition is more frequent than a stranded one. With prepositional verbs, i.e., verbs that license a (formal) preposition, pied-piping is more frequent in simple relative clauses but stranding is the slightly more frequent variant in complex ones: "if the preposition [. . .] or the PP [. . .] is lexically associated with the main verb then increasing complexity leads to a decrease in pied-piping" (Hoffmann 2011, p. 168f). According to Hoffmann, this is because stranding facilitates the integration of lexical verb–preposition structures.

Even though the effect size is very small, the SPRT reported as Experiment 1 here showed that a stranded preposition speeds up reading at the gap site of the displaced *wh*-element in relative clauses. This can be considered support for Hoffmann's argument that adjacency of the (formal) preposition and its licensing head, the verb, facilitates processing. However, a pied-piped preposition results in a small processing advantage at the relative pronoun, which can, as the previous effect, be explained in terms of prediction-based processing. This speed-up cancels out the prolonged reading times of the pied-piped variant at the gap site after the lexical verb for simple relative clauses. Similarly, the advantage of a stranded preposition at the gap site cancels out the slight disadvantage at the left boundary of the relative clause in the simple-clause condition. Thus, once the reaction times of the entire stimulus are taken into account, there is no significant difference in reading times between a pied-piped and a stranded preposition in the simple condition. In the complex-clause condition, however, there is a significant difference between preposition stranding and pied-piping in the cumulative reading times with pied-piping being the faster variant. Hence, even if stranding has a local effect of facilitating the integration of verb–preposition structures, as claimed by Hoffmann, it does not facilitate reading on the more global level of the whole clause. Obviously, the distance between the verb and the preposition in pied-piping across a complex relative clause results in a lower processing load than the distance between the preposition and its complement in stranding across complex clauses.

There is further misalignment of data both between the rating experiments and the reading task and, albeit to a lesser extent, between the rating experiments and the corpus data. Hoffmann's judgment task revealed a general preference for stranding in relative clauses, which is enhanced by an increase in structural complexity. A very similar picture is displayed in Experiment 2: Stranding is preferred over pied-piping, albeit in both complexity conditions. The variant with the better ratings—preposition stranding—is the variant that required more reading time in Experiment 1, at least in complex relative clauses. The variant with the lower ratings both in this and in Hoffmann's experiment—pied-piping—is the much more frequent one in Hoffmann's corpus analysis

There are several possible explanations for these discrepancies. The misalignment of corpus and SPRT data could be an effect of how structural complexity is operationalized. Hoffmann used a dichotomy based on Lu's (2002) complexity metric counting the number of chunks in a structure, whereas in the experiment here, COMPLEXITY was operationalized as the presence (or the absence) of a two-word bridging structure. It is possible that separating the preposition and the verb in a pied-piped construction causes processing difficulties if the distance is increased by more than two words. The corpus studies by Trotta (1998) and Gries (2002) point into this direction, because an increase in distance between the displaced element and the extraction results in a higher proportion of pied-piped prepositions. Thus, the results of Experiment 1 are not necessarily in conflict with the findings from the corpus studies. A second explanation is that even if preposition stranding is more difficult to read, it could still be easier to produce, as discussed by (Kunter 2017) for cognitive complexity of the comparative and the genitive alternation. Gries and Hoffmann make reference to Hawkins (1999) analysis of filler-gap structures, where inferring which filler is related to a gap imposes processing load on the hearer. Gries, however, also assumes that planning and production costs contribute to the cognitive load associated with the stranded variant. So, maybe speakers choose a variant that reduces cognitive complexity for them rather than the hearers, which is then reflected in the distribution of the variants in corpora.

Production vs. perception costs could also explain the differences between corpus findings and rating data. Stranding was the preferred option in both Hoffmann's magnitude estimation task and the split rating experiment reported here. In the corpus studies, pied-piping was the more frequent variant on the whole but in complex relative clauses, stranding was slightly more frequent in Hoffmann's analysis. If speakers make use of pied-piping to reduce cognitive complexity for themselves, pied-piping could still result in cognitive complexity on the hearer's end and hence be rated lower (see Radford et al. (2012) for a similar explanation for asymmetries between rating and corpus data on pruning and copying). Again, the operationalization of structural complexity could also play a role. There might be a cut-off point for the acceptability of preposition stranding for a longer distance between the preposition and the verb, i.e., pied-piping might be preferred with longer bridging structures. Trotta (1998, p. 207) presents some examples of unacceptable uses of stranding in the relative clause (albeit with non-complement PPs), stating that "preposition fronting becomes virtually obligatory if the distance between the preposition and its VP (or NP, AdjP, etc.) is too great".

Another potential explanation is genre or mode. Hoffmann pointed out a correlation between formal genres and a preference for pied-piping. Since the rating task operationalized acceptability as naturalness, the high preference for stranding, the informal variant, could as well be attributed to this difference. This begs the question of whether this is also the reason why the preferred variant in Experiment 2 is the one that resulted in the longest reading times in Experiment 1. However, a reading experiment cannot necessarily be considered a more formal context. First, the experimental setting was the same for both experiments, second, the experimental stimuli did not include overly formal English. The misalignment of data, in this case, cannot be explained by the operationalization of factors either, because the materials contained the exact same sentences. Rather, the type of task seems to play a role: Reading tracks processing (quasi-) online, whereas the sentences are compared and rated after processing is completed. In an offline task, the advantages of overt signals as to the function of constituents or which elements to expect next play a lesser role and may easily be overridden by other factors. This could explain why participants choose a form as the more acceptable option that is more difficult to read for them.

The question remains if the alignment between reading data and corpus data provides support for the role of cognitive complexity as a determinant of grammatical variation, or in other words, whether writers use pied-piping because this variant is the easier to process one at least with complex relative clauses. The answer is certainly not a clear yes. First, it has to be borne in mind that the role of prescriptivism cannot be factored out in the analysis of formal written language. Second, speakers, or writers in this context, might also benefit from the use of a pied-piped construction and their linguistic behavior might be driven by less altruistic factors. Third, to draw more general conclusions about preposition placement and processing, further syntactic contexts have to be tested.

6. Summary and Conclusions

This paper reported two experiments on variation between preposition stranding and pied-piping in English relative clauses. The aim was to shed light on the question of whether grammatical variation is determined by cognitive complexity, as often claimed in the pertinent literature. The review of the literature showed that there are quite different conceptions of what contributes to processing complexity. What is more, there is no consensus on which of the two variants is associated with a higher cognitive load and why. This is partly reflected in the data from the two experiments reported here. Not only do the results from the reading experiment and the rating task stand in contrast, but they also depart from what is expected on the basis of corpus data. Hence, this paper does not provide a clear answer either to the question of whether cognitive complexity is a determinant of preposition placement in English. What it does show, however, is that the mechanism behind grammatical variation might be more complex than the literature suggests. Even though the experiments here looked at preposition placement in a specific syntactic context only, they indicate that explanations based on Rohdenburg's Complexity Principle ("in the case of more or less explicit constructional options the more explicit one(s) will tend to be preferred in cognitively complex environments" (Rohdenburg 2003, p. 217)) or similar ideas might oversimplify matters. The question of which variant is the more difficult one is too general and should be narrowed down to at what point a certain variant induces difficulties and why. Claims that continuities in word order have clear benefits over discontinuous structures are much too coarse-grained. In addition, it is not entirely clear whether speakers avoid cognitively complex forms to relieve the hearer from cognitive effort or for their own benefit. This underlines the need to combine linguistic research methods to shed light on grammatical variation from different angles.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Linguistics and English Language Ethics Committee at the University of Edinburgh.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data management plan approved by the Ethics Committee stating that the data will not be made publicly available.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Critical Stimuli Experiment 1 and 2

- Chris provides the answer which Ashley searched for really desperately.
- Dwight mentions the lecture which Sue listened to very attentively.
- George describes the method which Rose approved of somewhat prematurely.
- Jack achieves the success which Sarah dreamed of rather regularly.
- Paul attends the event which Agnes trained for really ambitiously.
- Rob offers the support which Faye counted on truly despairingly.
- Dwayne outlines the danger which Beth escaped from somewhat unknowingly.
- John faces the problem which Anna coped with very effectively.

- Mike performs the miracle which Maureen prayed for very expectantly.
- Neal adopts the approach which Lyn argued for rather convincingly.
- Sam recalls the aspect which Jane agreed with fairly explicitly.
- Tom accepts the idea which Helen fought for highly aggressively.
- Bruce submits the report which Stella asked for only informally.
- Ken avoids the mistake which Susan paid for very considerably.
- Mark reflects the ordeal which Trish suffered from almost undoubtedly.
- Pete depicts the moment which Kate waited for slightly impatiently.
- Ted prefers the question which Jill began with rather hesitantly.
- Tim loathes the pressure which Alice died of almost literally.
- Bill enjoys the subject which Mary turned to very passionately.
- Craig issues the license which Liz applied for really successfully.
- Frank fancies the topic which Rachel wrote on somewhat exhaustively.
- Jim dislikes the matter which Sophie spoke of fairly excitedly.
- Max follows the debate which Meg prepared for fairly effectively.
- Nick welcomes the measure which Claire voted for only reluctantly.

Appendix B

Table A1. Summary of the first linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.6400	0.0437	197.927	0
PREVPC1	−0.4279	0.0167	−25.581	0
PREVPC2	0.0875	0.0259	3.384	0.000743
SURPRISALGAP	−0.0015	0.0025	1.235	0.217301
TRIAL	0.0031	0.0005	−1.927	0.054237
ORDER = stranded	−0.0155	0.0428	−0.362	0.717392
COMPLEXITY = simple	0.0204	0.0431	0.473	0.636577
TRIAL:ORDER = stranded	−0.0006	0.0007	−0.897	0.369750
TRIAL:COMPLEXITY = simple	−0.0007	0.0007	−0.953	0.340652
ORDER = stranded:COMPLEXITY = simple	−0.0326	0.0603	−0.540	0.589457
TRIAL:ORDER = stranded:COMPLEXITY = simple	0.0004	0.0010	0.382	0.702242

Table A2. Summary of the second linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.6450	0.0416	207.708	0
PREVPC1	−0.4278	0.0167	−25.575	0
PREVPC2	0.0877	0.0259	3.394	0.000716
SURPRISALGAP	0.0031	0.0025	1.223	0.221572
TRIAL	−0.0010	0.0004	−2.402	0.016509
ORDER = stranded	−0.0253	0.0343	−0.735	0.462438
COMPLEXITY = simple	0.0104	0.0343	0.302	0.762591
TRIAL:ORDER = stranded	−0.0004	0.0005	−0.883	0.377307
TRIAL:COMPLEXITY = simple	−0.0005	0.0005	−0.957	0.338961
ORDER = stranded:COMPLEXITY = simple	−0.0121	0.0277	−0.436	0.663077

Table A3. Summary of the third linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.6470	0.0413	209.220	0
PREVPC1	−0.4281	0.0167	−25.620	0
PREVPC2	0.0878	0.0258	3.396	0.000711
SURPRISALGAP	0.0031	0.0025	1.253	0.210682
TRIAL	−0.0010	0.0004	−2.392	0.016950
ORDER = stranded	−0.0310	0.0317	−0.976	0.329726
COMPLEXITY = simple	0.0039	0.0309	0.126	0.899421
TRIAL:ORDER = stranded	−0.0004	0.0005	−0.902	0.367061
TRIAL:COMPLEXITY = simple	−0.0005	0.0005	−0.945	0.345050

Table A4. Summary of the fourth linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.6600	0.0389	222.367	0
PREVPC1	−0.4274	0.0167	−25.565	0
PREVPC2	0.0866	0.0258	3.356	0.000821
SURPRISALGAP	0.0030	0.0025	1.222	0.222097
TRIAL	−0.0012	0.0004	−2.392	0.016950
ORDER = stranded	−0.0310	0.0004	−3.504	0.000480
COMPLEXITY = simple	0.0044	0.0309	0.141	0.887915
TRIAL:COMPLEXITY = simple	−0.0005	0.0005	−0.956	0.339556

Table A5. Summary of the fifth linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.6700	0.0374	231.941	0
PREVPC1	−0.4267	0.0167	−25.497	0
PREVPC2	0.0863	0.0258	3.347	0.000849
SURPRISALGAP	0.0032	0.0025	1.282	0.200111
TRIAL	−0.0015	0.0003	−2.392	0
ORDER = stranded	−0.0549	0.0174	−3.163	0.002560
COMPLEXITY = simple	−0.0212	0.0155	−1.364	0.179205

Table A6. Summary of the sixth linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.705387	0.025602	340.027	0
PREVPC1	−0.425647	0.016744	−25.420	0
PREVPC2	0.087772	0.025768	3.406	0.000686
TRIAL	−0.001484	0.000257	−5.774	0
ORDER = stranded	−0.04756	0.016467	−2.888	0.006025
COMPLEXITY = simple	−0.020340	0.015444	−1.317	0.194425

Table A7. Summary of the final linear-mixed effect model for the logged RTs in the spill-over region of Experiment 1, with a random intercept for SUBJECT and random Slopes for ORDER and COMPLEXITY over SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	8.6970	0.0246	353.034	0
PREVPC1	−0.4275	0.0168	−25.606	0
PREVPC2	0.0902	0.0257	3.504	0.000479
TRIAL	−0.0015	0.0003	−5.776	0
ORDER = stranded	−0.0323	0.0165	−2.894	0.005935

Table A8. Summary of first final linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	5.9680	0.3096	19.275	0
LOGINTENSIFIER	0.3642	0.0339	10.739	0
SURPRISALGAP	−0.0024	0.0045	−0.531	0.5964
TRIAL	−0.0041	0.0007	−5.885	0
ORDER = stranded	−0.0908	0.0617	−1.470	0.1419
COMPLEXITY = simple	0.1202	0.0613	−1.960	0.0502
TRIAL:ORDER = stranded	0.0014	0.0010	1.347	0.1784
TRIAL: COMPLEXITY = simple	0.0007	0.0010	0.688	0.4919
ORDER = stranded: COMPLEXITY = simple	0.1701	0.0877	1.939	0.0528
TRIAL: ORDER = stranded: COMPLEXITY = simple	−0.0024	0.0014	−1.649	0.0995

Table A9. Summary of the second linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	5.9440	0.3066	19.387	0
LOGINTENSIFIER	0.3640	0.0339	10.738	0
TRIAL	−0.0041	0.0007	−5.879	0
ORDER = stranded	−0.0959	0.0610	−1.574	0.1159
COMPLEXITY = simple	−0.1223	0.0612	−1.998	0.0459
TRIAL:ORDER = stranded	0.0013	0.0010	1.327	0.1848
TRIAL: COMPLEXITY = simple	0.0007	0.0967	0.699	0.4844
ORDER = stranded: COMPLEXITY = simple	0.1710	0.0962	1.950	0.0515
TRIAL: ORDER = stranded: COMPLEXITY = simple	−0.0024	0.0964	−1.640	0.1014

Table A10. Summary of the third linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	5.9460	0.3069	19.378	0
LOGINTENSIFIER	0.3603	0.0339	10.640	0
TRIAL	−0.0036	0.0006	−5.783	0
ORDER = stranded	−0.0333	4.7490	−0.701	0.484
COMPLEXITY = simple	−0.0608	−0.0608	−1.257	0.209
TRIAL:ORDER = stranded	0.0002	0.0007	0.236	0.813
TRIAL: COMPLEXITY = simple	−0.0004	0.0007	−0.605	0.545
ORDER = stranded: COMPLEXITY = simple	0.0432	0.0403	1.074	0.283

Table A11. Summary of the fourth linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	5.9410	0.3058	19.425	0
LOGINTENSIFIER	0.3605	0.0338	10.651	0
TRIAL	−0.0035	0.0005	−6.748	0
COMPLEXITY = simple	−0.0612	0.0484	−1.265	0.206
ORDER = stranded	−0.0243	0.0284	−0.855	0.393
TRIAL: COMPLEXITY = simple	−0.0004	0.0007	−0.599	0.549
ORDER = stranded: COMPLEXITY = simple	0.0435	0.0402	1.082	0.280

Table A12. Summary of the fifth linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	5.9520	0.3052	19.503	0
LOGINTENSIFIER	0.3605	0.0338	10.654	0
TRIAL	−0.0037	0.0004	−9.873	0
ORDER = stranded	−0.0246	0.0284	−0.867	0.38597
COMPLEXITY = simple	−0.0846	0.0284	−2.978	0.00298
ORDER = stranded: COMPLEXITY = simple	0.0441	0.0402	1.097	0.27296

Table A13. Summary of the sixth linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random intercepts for ITEM and SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	5.9500	0.3052	19.494	0
LOGINTENSIFIER	0.3595	0.0338	10.627	0
TRIAL	−0.0037	0.0004	−9.873	0
ORDER = stranded	−0.0026	0.0201	−0.131	0.89578
COMPLEXITY = simple	−0.0627	0.0202	−3.107	0.00195

Table A14. Summary of the final linear-mixed effect model for the logged RTs in the wrap-up region of Experiment 1, with random interecepts for ITEM and SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	5.9470	0.3040	19.564	0
LOGINTENSIFIER	0.3596	0.0338	10.652	0
COMPLEXITY = simple	−0.0626	0.0202	−3.108	0.00194
TRIAL	−0.0037	0.0004	−9.885	0

Table A15. Summary of the first linear-mixed effect model for the logged RTs at the relative pronoun of Experiment 1, with a random intercept for SUBJECT and a random slope for ORDER over SUBJECT.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	8.5480	0.0687	124.482	0
PREVRELPC1	0.4022	0.0174	23.099	0
PREVRELPC2	−0.0752	0.0286	−2.627	0.008744
SURPRISALREL	−0.0026	0.0078	−0.338	0.735507
TRIAL	−0.0014	0.0004	−3.816	0.000144
ORDER = stranded	0.0547	0.0329	1.664	0.096967
TRIAL:ORDER = stranded	−0.0003	0.0005	−0.575	0.565480

Table A16. Summary of the second linear-mixed effect model for the logged RTs at the relative pronoun of Experiment 1, with a random intercept for SUBJECT and a random slope for ORDER over SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	8.5280	0.0343	248.549	0
PREVRELPC1	0.4025	0.0174	23.145	0
PREVRELPC2	−0.0752	0.0286	−2.628	0.008722
TRIAL	−0.0014	0.0004	−3.816	0.000144
ORDER = stranded	0.0564	0.0325	1.734	0.083803
TRIAL:ORDER = stranded	−0.0003	0.0005	−0.571	0.568090

Table A17. Summary of the final linear-mixed effect model for the logged RTs at the relative pronoun of Experiment 1, with a random intercept for SUBJECT and a random slope for ORDER over SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	8.5350	0.0317	269.547	0
PREVRELPC1	0.4027	0.0174	23.168	0
PREVRELPC2	−0.0748	0.0286	−2.616	0.00904
TRIAL	−0.0015	0.0003	−5.630	0
ORDER = stranded	0.0406	0.0171	2.376	0.02258

Table A18. Summary of the first linear-mixed effect model for the logged cumulative RT in the simple condition of Experiment 1, with a random intercept for SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.0700	0.0973	124.135	0
SURPRISALGAP	0.0021	0.0030	0.691	0.490
SURPRISALREL	−0.0020	0.0084	−0.237	0.813
TRIAL	−0.0033	0.0004	−8.449	0
ORDER = stranded	−0.0070	0.0348	−0.202	0.840
TRIAL:ORDER = stranded	0.0002	0.0006	0.309	0.757

Table A19. Summary of the second linear-mixed effect model for the logged cumulative RT in the simple condition of Experiment 1, with a random intercept for SUBJECT.

	Estimate	Error	t-Value	*p*-Value
Intercept	12.0600	0.0777	155.251	0
SURPRISALGAP	0.0020	0.0029	0.665	0.506
TRIAL	−0.0033	0.0004	−8.459	0
ORDER = stranded	−0.0057	0.0343	−0.166	0.868
TRIAL:ORDER = stranded	0.0002	0.0006	0.316	0.752

Table A20. Summary of the third linear-mixed effect model for the logged cumulative RT in the simple condition of Experiment 1, with a random intercept for SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.0500	0.0760	158.520	0
SURPRISALGAP	0.0020	0.0029	0.673	0.501
TRIAL	−0.0032	0.0003	−11.576	0
ORDER = stranded	0.0039	0.0162	0.240	0.810

Table A21. Summary of the third linear-mixed effect model for the logged cumulative RT in the simple condition of Experiment 1, with a random intercept for SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.0500	0.0759	158.826	0
SurprisalGap	0.0022	0.0028	0.806	0.421
Trial	−0.0032	0.0003	−11.585	0

Table A22. Summary of the final linear-mixed effect model for the logged cumulative RT in the simple condition of Experiment 1, with a random intercept for SUBJECT.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.0800	0.0672	179.77	0
TRIAL	−0.0032	0.0003	−11.73	0

Table A23. Summary of the first linear-mixed effect model for the logged cumulative RT in the complex condition of Experiment 1, with random intercepts for SUBJECT and ITEM.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.2400	0.1099	111.354	0
SURPRISALGAP	0.0000	0.0037	−0.006	0.995
SURPRISALREL	0.0122	0.0105	1.157	0.249
TRIAL	−0.0033	0.0004	−8.522	0
ORDER = stranded	0.0360	0.0356	1.012	0.312
TRIAL:ORDER = stranded	0.0003	0.0006	0.502	0.616

Table A24. Summary of the second linear-mixed effect model for the logged cumulative RT in the complex condition of Experiment 1, with random intercepts for SUBJECT and ITEM.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.2400	0.1031	118.752	0
SURPRISALREL	0.0121	0.0105	1.154	0.250
TRIAL	−0.0033	0.0004	−8.529	0
ORDER = stranded	0.0359	0.0345	1.040	0.299
TRIAL:ORDER = stranded	0.0003	0.0006	0.502	0.616

Table A25. Summary of the third linear-mixed effect model for the logged cumulative RT in the complex condition of Experiment 1, with random intercepts for SUBJECT and ITEM.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.2300	0.1021	119.794	0
SURPRISALREL	0.0120	0.0105	1.148	0.25250
TRIAL	−0.0031	0.0003	−11.376	0
ORDER = stranded	0.0510	0.0171	2.985	0.00298

Table A26. Summary of the final linear-mixed effect model for the logged cumulative RT in the complex condition of Experiment 1, with random intercepts for SUBJECT and ITEM.

	Estimate	Error	*t*-Value	*p*-Value
Intercept	12.323241	0.064500	191.057	0
TRIAL	−0.003145	0.000277	−11.356	0
ORDER = stranded	0.042534	0.015437	2.755	0.0061

Table A27. Summary of the first linear-mixed effect model for Experiment 2, with random intercepts for Subject and Item and random slopes for Order over Subject and over Item.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	35.79	2.514	14.239	0
Order = stranded	12.28	3.910	3.142	0.00199
Trial	0.02206	0.03091	0.714	0.47560
Complexity = simple	6.693	2.742	2.441	0.01483
Trial: Order = stranded	0.04504	0.04493	1.002	0.31637
Trial:Complexity = simple	−0.06689	0.0437	−1.507	0.13199
Order = stranded:Complexity = simple	−2.957	3.910	−0.756	0.44967
Trial:Order = stranded:Complexity = simple	0.005289	0.06406	0.008	0.99341

Table A28. Summary of the second linear-mixed effect model for Experiment 2, with random intercepts for Subject and Item and random slopes for Order over Subject and over Item.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	35.79896	2.37705	15.060	0
Order = stranded	12.27011	3.52172	3.484	0.000705
Trial	0.02194	0.02700	0.813	0.416622
Complexity = simple	6.67930	2.15845	3.094	0.002024
Trial: Order = stranded	0.04530	0.03176	1.426	0.154091
Trial:Complexity = simple	−0.06664	0.03199	−2.083	0.037490
Order = stranded:Complexity = simple	−2.92800	1.77205	−1.652	0.098777

Table A29. Summary of third linear-mixed effect model for Experiment 2, with random intercepts for Subject and Item and random slopes for Order over Subject and over Item.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	34.63576	2.23243	15.515	0
Order = stranded	14.68881	3.09013	4.753	0.00001
Trial	0.04335	0.02244	1.932	0.05363
Complexity = simple	6.60669	2.15851	3.061	0.00226
Trial:Complexity = simple	−0.06597	0.03200	−2.062	0.03947
Order = stranded:Complexity = simple	−2.83998	1.77160	−1.603	0.10923

Table A30. Summary of final linear-mixed effect model for Experiment 2, with random intercepts for Subject and Item and random slopes for Order over Subject and over Item.

	Estimate	**Error**	***t*-Value**	***p*-Value**
Intercept	35.37642	2.18449	16.194	0
Order = stranded	13.26914	2.95968	4.483	0.00003
Trial	0.04271	0.02245	1.902	0.05744
Complexity = simple	5.13081	1.95504	2.624	0.00881
Trial:Complexity = simple	−0.06484	0.03201	−2.026	0.04306

References

Baayen, R. Harald, and Petar Milin. 2010. Analyzing reaction times. *International Journal of Psychological Research* 3: 12–28. [CrossRef]

Baayen, R. Harald, Davidson Douglas, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412. [CrossRef]

Bartek, Brian, Richard L. Lewis, Shravan Vasishth, and Mason R. Smith. 2001. In Search of On-line Locality Effects in Sentence Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37: 1178–98. [CrossRef] [PubMed]

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

Berlage, Eva. 2009. Prepositions and postpositions. In *One Language, Two Grammars? Differences between British and American English*. Edited by Günter Rohdenburg and Julia Schlüter. Cambridge: Cambridge University Press, pp. 130–48.

Berlage, Eva. 2014. *Noun Phrase Complexity in English (Studies in English Language)*. Cambridge: Cambridge University Press.
Bresnan, Joan, and Marilyn Ford. 2010. Predicting Syntax: Processing Dative Constructions in American and Australian Varieties of English. *Language* 86: 186–213. [CrossRef]
Gries, Stefan Th. 2002. Preposition stranding in English: Predicting speakers' behaviour. In *Proceedings of the Western Conference on Linguistics*. Edited by Vida Samiian. Fresno: California State University, vol. 12, pp. 230–41.
Gries, Stefan Th. 2003. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. London and New York: Continuum Press.
Günther, Christine. 2018. A difficult to explain phenomenon—Increasing complexity in the prenominal position. *English Language & Linguistics*, 1–26. [CrossRef]
Hawkins, John. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75: 245–85. [CrossRef]
Hawkins, John. 2000. The relative order of prepositional phrases in English: Going beyond manner-place-time. *Language Variation and Change* 11: 231–66. [CrossRef]
Hawkins, John. 2003. Why are zero-marked phrases close to their heads? In *Determinants of Grammatical Variation in English*. Edited by Günter Rohdenburg and Britta Mondorf. Berlin and New York: Mouton de Gruyter, pp. 175–204.
Hawkins, John. 2014. *Cross-linguistic Variation and Efficiency*. Oxford: Oxford University Press.
Hoffmann, Thomas. 2011. *Preposition placement in English*. Cambridge: Cambridge University Press.
Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.
Just, Marcel Adam, and Patricia Carpenter. 1980. A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review* 87: 329–54. [CrossRef] [PubMed]
Konieczny, Lars. 2000. Locality and Parsing Complexity. *Journal of Psycholinguistic Research* 29: 627–45. [CrossRef] [PubMed]
Kunter, Gero. 2017. *Processing Complexity and the Alternation between Analytic and Synthetic Forms in English*. Habilitationsschrift: Heinrich-Heine-Universität Düsseldorf.
Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106: 1126–77. [CrossRef] [PubMed]
Lohse, Barbara, John Hawkins, and Thomas Wasow. 2004. Domain minimization in English verb-particle constructions. *Language* 80: 238–61. [CrossRef]
Lu, Bingfu. 2002. How does language encode performance limitation into its structure. Unpublished manuscript.
Mitchell, Don C., and David W. Green. 1978. The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology* 30: 609–36. [CrossRef]
Pullum, Geoffrey, and Rodney Huddleston. 2002. Prepositions and preposition phrases. In *Huddleston, Rodney & Geoffrey Pullum. The Cambridge Grammar of the English Language*. Cambridge: CUP, pp. 597–661.
Radford, Andrew, Claudia Felser, and Oliver Boxwell. 2012. Preposition copying and pruning in present-day English. *English Language & Linguistics* 16: 403–26.
Rohdenburg, Günter. 2003. Cognitive complexity and horror aequi as factors determining the use of interrogative clause linkers in English. In *Determinants of Grammatical Variation in English*. Edited by Rohdenburg Günter and Britta Mondorf. Berlin and New York: Mouton de Gruyter, pp. 205–49.
Ross, John R. 1967. Constraints on Variables in Syntax. Ph.D. dissertation, MIT, Cambridge, MA, USA.
Rühlemann, Christoph, and Stefan Th. Gries. 2020. Speakers advance-project turn completion by slowing down: A multifactorial corpus analysis. *Journal of Phonetics* 80: 100976. [CrossRef]
Trotta, Joe. 1998. *Wh*-Clauses in English: Aspects of Theory and Description. Ph.D. thesis, Göteborg Universitet, Gothenburg, Sweden.
Vasishth, Shravan, and Heiner Drenhaus. 2011. Locality in German. *Dialogue and Discourse* 1: 59–82. [CrossRef]
Wasow, Thomas, and Jennifer Arnold. 2003. Post-verbal Constituent Ordering in English. In *Determinants of Grammatical Variation in English*. Edited by Günter Rohdenburg and Britta Mondorf. Berlin and New York: Mouton de Gruyter, pp. 119–54.