

Sina Bosch, Ilaria De Cesare, Claudia Felser & Ulrike Demske

2022

A Multi-methodological Approach to Word Order Variation in  
German Infinitival Complementation



in

Robin Hörnig, Sophie von Wietersheim, Andreas Konietzko & Sam Featherston (eds.),  
*Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data*

pp. 281–298

Tübingen: University of Tübingen

<https://publikationen.uni-tuebingen.de/xmlui/handle/10900/119301>



# A Multi-methodological Approach to Word Order Variation in German Infinitival Complementation

Sina Bosch, Ilaria De Cesare, Claudia Felser & Ulrike Demske

University of Potsdam

*bosch/decesare/felser/udemske@uni-potsdam.de*

## 1 Introduction

German exhibits substantial variability regarding the linearization of infinitive-embedding verbs and their complements: Infinitival complements can either be extraposed to the right of the matrix verb (1), intraposed to its left (2), or form a discontinuous infinitival construction, the so-called 'Third Construction' (3) (Haider, 2010).

- (1) EXTRAPOSITION  
*dass Fred versucht [den Kuchen zu schneiden]*  
that Fred tries the cake to cut
- (2) INTRAPOSITION  
*dass Fred [den Kuchen zu schneiden] versucht*  
that Fred the cake to cut tries
- (3) THIRD CONSTRUCTION  
*dass Fred [den Kuchen] versucht [zu schneiden]*  
that Fred the cake tries to cut  
'That Fred tries to cut the cake.'

These syntactic variation patterns have been analysed extensively in formal linguistic theory (see Haider, 2010; Reis, 2001; Wurmbrand, 2001 among others). However, the focus of our present investigation will lie on the question of which factors limit word order variation in German infinitival complementation and determine word order preferences during actual language performance. By approaching language performance from a multi-methodological perspective, we assess, firstly, whether frequency distributions of different word order variants determine speakers' preference patterns as predicted by usage-based approaches to language performance (e.g. Bybee, 2006; Bybee & Beckner, 2010). Secondly, factors other than frequency may also play a role in determining speakers' choices, such as the ease of producing or comprehending a particular word order variant. Therefore, we also examine whether and to what extent processing economy constraints (e.g. Hawkins, 1994; Gibson, 2000) might influence speakers' preferences in syntactic variation contexts.

To examine performance patterns in word order variation in German infinitival complementation, we draw together results from four different data sources including (i) a corpus study to establish the frequency distributions of the different infinitival word order patterns, (ii) a spoken production experiment, (iii) two scalar acceptability judgement experiments, which will allow us to evaluate speakers' word order preferences against the predictions of usage-based approaches to linguistic variation, and (iv) a reading-time experiment to assess the influence of processing economy constraints on speakers' preference patterns.

## 2 Limitations on Word Order Variation in German Infinitival Complementation

### 2.1 Usage-based Approaches

The present study focuses on two potential sources of limitations on word order variation in German infinitival complementation. Usage-based approaches to language processing and representation argue that frequency of occurrence determines speakers' preference patterns in variation contexts. Accordingly, both language use and grammatical representation are regarded as emergent, i.e. not given a priori, but coming about through the frequency of input and the way the brain responds to the experience of language use. Hence, implicit procedural knowledge about input probabilities and statistical patterns of linguistic expressions are seen as the basis on which a speaker's linguistic knowledge and use are built (e.g. Larsen-Freeman, 1997; Bybee & Hopper, 2001; Gahl & Garnsey, 2004; Bybee, 2006; Bybee & Beckner, 2010). In the context of syntactic variation in German infinitival complementation, such an approach to language representation would predict that language users' preferences in the choice of word order variants – both in *production* and in *comprehension* – should mirror their frequency distributions across the language.

A number of previous studies have examined the predictions of usage-based accounts by assessing the relationship between frequency distributions as measured by synchronous corpus probabilities and speakers' preference patterns with respect to different linguistic phenomena in experimental tasks, such as acceptability judgements (e.g. Radford et al., 2012; Schmid et al., 2005). Taken together, many of these studies revealed imperfect correlations between frequency distributions and acceptability judgements, such that highly frequent alternates received high acceptability ratings, while low-frequency variants which are rarely attested in corpora might nevertheless be deemed acceptable (Bader & Häussler, 2010; Featherston, 2005; Kempen & Harbusch, 2005; Newmeyer, 2003). As regards the phenomenon under investigation in the present study, Bayer et al. (2005) also conducted a multi-methodological study combining a systematic corpus search with several experimental tasks. The authors found that, although intraposed infinitival complements only occurred infrequently in the corpora they examined, speakers' judgement patterns showed mean ratings for intraposed structures in the lower intermediate range (i.e. a rating of 2.1 on a 5-point Likert scale). The present study tests word order variation in German infinitival complementation for its correlation with frequency distributions in a more fine-grained manner by (i) conducting an in-depth corpus study to obtain frequency distributions of the different word order variants in Present-Day German, which will then serve as a baseline against which speakers' preferred choices will be evaluated, and (ii) by restricting the investigation of infinitive-embedding verbs to a small sample set of subject control verbs and assessing their word order preferences systematically across all three word order patterns, Extraposition, Intraposition, and Third Construction.

### 2.2 Processing Economy Constraints

Notions like complexity, economy, efficiency, and processing ease have long been argued to play a major role in explaining language processing and representation. Following Chomsky (2005), for example, principles of efficient computation play a crucial role in language representation and development just as they do in other cognitive domains. In psycholinguistic research, several processing models and principles have been proposed to capture the sources of processing cost in syntactic variation contexts. One of the most influential processing-based constraints goes back to Prideaux & Baker's (1986) principle of *Closure*, which favours word order patterns that allow for clausal units to be completed as soon as possible during language processing. In a similar fashion, Hawkins' (1994) principle of *Early Immediate Constituents*

claims that “words and constituents occur in the orders they do so that syntactic groupings and their immediate constituents (ICs) can be recognized (and produced) as rapidly and efficiently as possible in language performance” (p. 57). Alternative orderings of individual syntactic elements might thus result in more or less efficient constituent recognition. The *Dependency Locality Theory* proposed by Gibson (1998, 2000) postulates that a greater distance between dependent syntactic elements leads to higher structural integration costs. Taken together, all of these proposals share the assumption that nested or centre-embedded structures require more processing effort compared to non-nested structures.

With respect to the different word order patterns in German infinitival complementation, the only word order variant that fully avoids centre-embedded structures and does not break up any structural dependency is Extraposition, where the infinitival complement appears at the right clausal periphery (see example (1), repeated here as (4)).

- (4) EXTRAPOSITION  
*dass Fred versucht [den Kuchen zu schneiden]*  
 that Fred tries      the cake      to cut

In extraposed infinitival structures, the integrity of the infinitival complement is preserved and the dependency between the subject *Fred* in the superordinate clause and the head of its predicate *versucht* is not interrupted by an intervening phrasal or clausal constituent, as is the case for Intraposition structures. Extraposed infinitival constructions are also unlikely to trigger any initial structural or semantic misanalysis during language comprehension, in contrast to Third Constructions (see below for details). Accordingly, following Hawkins' (1994) and Gibson's (2000) principles of processing economy, extraposed infinitival complements are expected to produce the smallest processing cost relative to the other two word order variants. This is predicted despite the fact that German subordinate clauses, including complement clauses introduced by *dass* ('that'), normally show verb-final order.

The situation is different for intraposed infinitival complements (see example (2), repeated here as (5)).

- (5) INTRAPOSITION  
*dass Fred [den Kuchen zu schneiden] versucht*  
 that Fred the cake      to cut      tries

In intraposed infinitival constructions, the infinitival complement is centre-embedded between the subject *Fred* and the subcategorizing verb *versucht*, and thus the subject-verb dependency is broken up. Intraposition structures are also ambiguous between a mono- and a biclausal construal, as the embedded non-finite verb *schneiden* can either project an independent clausal domain (i.e. a CP) resulting in a biclausal structure, or the matrix predicate *versucht* and the non-finite verb *schneiden* can form a verbal complex within a monoclausal domain (Bech, 1955; Haider, 2010). Consequently, intraposed infinitival constructions are expected to elicit larger processing costs relative to extraposed structures.

Finally, the Third Construction represents a discontinuous construction breaking up two structural dependencies: Similar to intraposed infinitival complements, the distance between the subject *Fred* and the matrix predicate *versucht* is increased by placing the accusative noun phrase *den Kuchen* between those two elements. At the same time, the verbal phrase *zu schneiden* remains in an extraposed position to the right of the matrix predicate, breaking up the infinitival complement's structural integrity (see example (3), repeated here as (6)).

- (6) THIRD CONSTRUCTION  
*dass Fred [den Kuchen] versucht [zu schneiden]*  
 that Fred the cake      tries      to cut

Due to this discontinuous construal, the Third Construction gives rise to local ambiguity. From the perspective of incremental left-to-right processing, the noun phrase *den Kuchen* is likely to be initially analysed as the direct object of the infinitive-embedding verb *versucht*. Such a structural analysis would not only be the computationally simplest option based on the current linguistic input, but also allows for a thematic role to be assigned to the accusative noun phrase immediately. However, encountering the infinitive *zu schneiden* renders this initial analysis and interpretation untenable and forces the comprehender to backtrack and revise their analysis. Third Constructions may thus be comparable to ‘garden-path’ sentences, which have been studied extensively in sentence processing research (e.g. Frazier & Rayner, 1982). Revising the initially computed analysis and interpretation of Third Construction sentences might lead to a measurable increase in processing cost relative to both intraposed and extraposed infinitival complements. Here we will test whether the above processing economy considerations can predict the relative comprehension difficulty of the three word order variants under investigation by carrying out a reading-time study.

### 3 The Present Study

The present study uses a multi-methodological approach to word order variation in German infinitival complementation. We provide a synthesis of corpus-based and experimental investigations in order to test the predictability of language performance across different tasks from both (i) corpus frequency distributions and (ii) processing economy constraints. We first take a corpus-based perspective on German infinitival complementation to assess how the different word order patterns are distributed across different types of synchronous German corpora. The frequency distributions obtained from a systematic corpus analysis will provide the baseline against which the predictions from usage-based frameworks of language representation can be evaluated. We evaluate these predictions both for speakers’ *production* patterns obtained via an elicited spoken production task and for speakers’ *judgement* patterns as investigated in two scalar acceptability rating tasks. Secondly, the present study takes a processing perspective on syntactic variation patterns. In order to investigate to what extent processing economy might affect speakers’ word order preferences in German infinitival complementation, we administered a reading-time study to measure the relative speed at which different word order variants are processed. The combined findings of our corpus-based and experimental investigations will provide a more comprehensive and nuanced picture of the limits to grammatical variation in language performance and representation.

To be able to systematically analyse German speakers’ word order preferences for their correlation with frequency distributions and processing economy constraints, in the present study we restrict the number of infinitive-embedding verbs to a small sample of subject control verbs. Data collection across all of our present investigations focused on the same four verbs: *bedauern* (‘regret’), *ankündigen* (‘announce’), *beschließen* (‘decide’), and *versuchen* (‘try’). All of these verbs share the same control properties and argument structure but differ in their semantics, such that *ankündigen* and *bedauern* are factive/propositional predicates, whereas *beschließen* and *versuchen* are irrealis predicates (Wurmbrand, 2001). In addition, in Wurmbrand (2014) and Ritter & Wiltschko (2014) *versuchen* is characterized as simultaneous non-propositional verb and thus belongs to a class of infinitive-embedding verbs with differing temporal-aspectual properties: In particular, Grano (2011) and Brandner (2020) argue that the structural size of the infinitival complement selected by *versuchen* is not the same as with the other control verbs. Instead, the matrix verb *versuchen* and its embedded infinitive do not describe separate event times, but include a temporal extension which represents a subpart of the event as a whole. Although we are aware of the fact that semantic and structural factors,

notably the size of the embedded infinitive, might play a role as well, we do not focus on these factors in the present study (see Bosch et al., in press, for some discussion of these factors).

### 3.1 Corpus Study

In our corpus analysis, frequencies were drawn from a corpus including both written (tokens = 405,331,744) and spoken German (tokens = 2,360,000). The written sources were searched in the context of a related study (Bosch et al., in press) and ranged from newspaper articles to narratives and Twitter posts. In addition to this, for the present study we extracted sentences with infinitival complements from the Tüba-D/S and the Folk (Forschungs- u. Lehrkorpus für gesprochenes Deutsch) corpora, two collections of spontaneous speech data in both formal and informal contexts. However, no combination of the control verbs *ankündigen*, *bedauern*, and *beschließen* plus infinitival complement was attested in the spoken data. Only for the verb *versuchen* did we obtain relevant corpus hits. The total distribution of word order patterns across all corpora is summarised in Table 1.

**Table 1.** Corpus frequencies for the four subject control verbs (percentages in parentheses)

	<i>ankündigen</i>	<i>bedauern</i>	<i>beschließen</i>	<i>versuchen</i>
<b>Extraposition</b>	127 (100)	39 (97.5)	277 (97.9)	2286 (74.5)
<b>Intraposition</b>	0 (0)	1 (2.5)	6 (2.1)	676 (22)
<b>Third Construction</b>	0 (0)	0 (0)	0 (0)	106 (3.5)

As was expected, Extraposition was found to be the most frequent pattern for all verbs. However, whereas Intraposition and the Third Construction are not or only rarely attested for the verbs *bedauern*, *ankündigen* and *beschließen*, the verb *versuchen* shows more variation such that Intraposition is attested in 22 % and the Third Construction in 3.5 % of all cases (see Table 1). Interestingly, however, the Third Construction instances are found almost exclusively in the spoken ( $N = 41/106$ ) and in the Twitter corpus ( $N = 65/106$ ), whereas intraposed infinitival complements are practically absent from the spoken records ( $N = 1/676$ ). A Fisher's Exact Test on the absolute CORPUS COUNTS revealed that verbs and the distribution of attested word order patterns were significantly associated ( $p < .001$ ). Post hoc comparisons by pairs of verbs and attested word order patterns confirmed that the verbs *bedauern*, *ankündigen* and *beschließen* did not differ significantly with respect to the frequency distributions of the word order variants they occur with, but that *versuchen* differs significantly from each of the other verbs (vs *bedauern*:  $p < .01$ , vs *ankündigen*:  $p < .001$ , vs *beschließen*:  $p < .001$ ). Whether and to what extent these frequency distributions of word order variants in German infinitival complementation are also reflected in behavioural preference patterns will be investigated in the following sections.

### 3.2 Experimental Investigations on Frequency vs. Language Performance

A spoken production experiment as well as two acceptability judgement tasks examined whether and to what extent the above corpus frequencies are in accordance with speakers' preferences in language performance. Frequency-based approaches to language representation and use predict that for each verb, performance patterns should mirror the relative occurrence frequencies of the structural variants under investigation. However, previous research reported a non-linear relation, particularly between corpus probabilities and acceptability ratings: Highly

acceptable alternatives predominate corpus frequencies, whereas less optimal candidates – although potentially felicitous choices – only occur scarcely (Featherston, 2005; Bader & Häussler, 2010). We tested these predictions for German infinitival complementation by comparing results across different data sources.

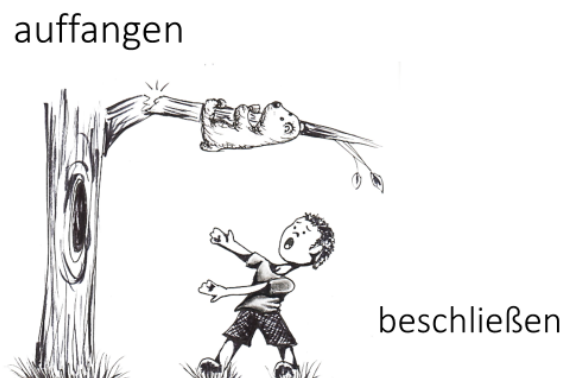
### 3.2.1 Experiment 1: Spoken Production

#### *Participants*

The spoken production experiment elicited speech data from 44 participants (12 male, 32 female) who were aged between 18 and 35 years (mean: 24.27 years). All participants were German native speakers living in Germany at the time of testing. The majority of them were students at different universities across the country who were recruited online, either via the SONA participant pool provided by the university of Potsdam or via e-mail contact. After completing the study, participants received 4€ as reimbursement for their participation.

#### *Materials*

Experimental materials consisted of picture cards depicting easily describable actions along with one of the four chosen infinitive-embedding control verbs and another content verb which helped participants to grasp the meaning of the depicted action (see Figure 1).



**Figure 1.** Example picture card for Experiment 1 with the control verb *beschließen* and the content verb *auffangen* ('to catch')

Participants were asked to look closely at the given picture and describe the action depicted in a single sentence. To make sure that infinitival complements preserved SOV structure, we instructed participants to place their picture descriptions in subordinate clauses. Hence, they produced sentences beginning with *Ich glaube, dass...* ('I think that...') and made use of both the infinitive-embedding control verb and the content verb presented next to the picture. However, they could freely choose the word order in which they arranged the control verb and the infinitive, thus allowing them to choose between Intraposition, Extraposition and Third Construction complements. For example, for Figure 1, we expected responses such as *Ich glaube, dass der Junge den Bären aufzufangen beschließt.* ('I think that the boy decides to catch the bear.'). or the corresponding Extraposition or Third Construction pattern. For each of the four control verbs, two picture cards with different depicted actions and thus, with different content verbs, were drawn. In addition, the experiment included 16 filler cards which only presented a content verb (e.g. *essen* 'eat') and a noun helping to describe the picture (e.g. *Suppe* 'soup'). These were not intended to elicit infinitival complements but only simple sentences such as *Ich glaube, dass das Kind die Suppe isst.* ('I think that the child will eat the soup'). The



full experimental design thus included 24 pictures. Experimental and filler picture cards were presented in pseudo-randomized order with no more than two critical items in a row. Complete lists of the experimental materials for Experiments 1-3 are available at the Open Science Framework website at <https://osf.io/kpjt4/>.

### Procedure

The spoken production experiment was administered in a quiet room on the University of Potsdam campus and consisted of two parts. Participants were first asked to answer some biographical questions and to give their voluntary consent to participating in the study. After reading the instructions, which explained the procedure of the experiment, participants were presented with three practice items to familiarize themselves with the experimental task before the actual experiment started. The set of picture cards was stacked face down on the table in front of the participants who were asked to take and describe one picture card after the other. Participants could take as much time as they needed for each picture. After finishing the description of a picture card, participants could continue with the next card at their own pace. The produced sentences were recorded with the help of a dictation machine and transcribed afterwards by a student assistant. On average, the elicited production experiment took approximately 20 minutes to complete.

### Results

Responses in which participants did not produce any infinitival complements in combination with one of the critical infinitive-embedding verbs were excluded from further analysis. This was the case for three sentences across the whole dataset. Counts and proportions (in parentheses) of the produced types of infinitival complements are shown in Table 2.

**Table 2.** Number of produced construction types (percentages in parentheses) in Experiment 1

	<i>ankündigen</i>	<i>bedauern</i>	<i>beschließen</i>	<i>versuchen</i>
<b>Extraposition</b>	74 (94.87)	66 (97.06)	70 (95.89)	60 (71.43)
<b>Intrapolition</b>	2 (2.56)	1 (1.47)	1 (1.37)	16 (19.05)
<b>Third Construction</b>	2 (2.56)	1 (1.47)	2 (2.74)	8 (9.52)

Extraposition was clearly the preferred word order pattern across all four matrix verbs. In analogy to the corpus data, Intrapolition and the Third Construction were hardly produced at all for the verbs *ankündigen*, *bedauern*, and *beschließen*. Only the verb *versuchen* elicited more variability in the linearization of its infinitival complements, such that intraposed infinitives were produced in 19.05 % of the cases and Third Constructions appeared in 9.52 % of all cases in the dataset. Generalized linear mixed-effect models on produced constructions across the four critical verbs were conducted using R (Baayen, 2008; R Core Team, 2017) and confirmed these numerical trends: *versuchen* elicited significantly fewer extraposed infinitival complements relative to all other subject control verbs (vs. *beschließen*:  $b = 2.456$ ;  $t = 3.601$ ;  $p < .001$ ; vs. *bedauern*:  $b = 2.833$ ;  $t = 3.565$ ;  $p < .001$ ; vs. *ankündigen*:  $b = 2.243$ ;  $t = 3.650$ ;  $p < .001$ ), but instead significantly more intraposed infinitives as well as Third Constructions. The other three control verbs, however, did not statistically differ (*beschließen* vs. *bedauern*:  $b = 0.377$ ;  $t = 0.397$ ;  $p = .692$ ; *beschließen* vs. *ankündigen*:  $b = -0.213$ ;  $t = -0.264$ ;  $p = .792$ ; *bedauern* vs. *ankündigen*:  $b = -0.589$ ;  $t = -0.654$ ;  $p = .513$ ).

Taken together, the data from our spoken production experiment and from our corpus study on written and spoken German yield parallel findings: (i) Extraposed infinitival complements are clearly preferred over the other two word order options across all infinitive-embedding control verbs, and (ii) the verb *versuchen* exhibits the highest degree of variability across the three word order variants. Hence, speakers' word order preferences in our spoken production data match corresponding corpus frequency distributions.

### 3.2.2 Experiment 2: Acceptability Judgements

In order to investigate whether and to what extent the observed word order preferences in corpus and production data extend to the domain of language comprehension, we conducted two complementary acceptability judgement experiments. Experiment 2A examines the acceptability of extraposed infinitival complements and of the Third Construction, and replicates an experimental study reported in Bosch et al. (in press) with another participant group and with a reduced set of experimental stimuli. Experiment 2B, which was also conducted and reported in Bosch et al. (in press), was complementary to the former and examined the acceptability of extraposed versus intraposed infinitival complements. Here we will present the combined findings of the two experiments relating to the acceptability patterns of the three word order variants Extraposition, Intraposition, and Third Construction.

#### *Participants*

Experiment 2A included 37 German adult native speakers (5 male, 32 female) aged between 18 and 32 years (mean age: 23.24 years). Experiment 2B tested 56 German adult native speakers (22 male, 34 female) aged between 20 and 42 years (mean age: 32.6 years). All participants were living in Germany at the time of testing and were recruited among the student and working communities in and around Potsdam and/or Berlin, either via the SONA participant pool provided by the university of Potsdam or via e-mail contact. None of them had any background in linguistics or were informed about the ultimate purpose of the study before testing, and all of them received 4€ as a small reimbursement for their participation.

#### *Materials*

The stimulus materials for Experiments 2A and 2B consisted of the four chosen verbs and infinitival complements in different word order variants. To avoid using verb-second (i.e. main clause) structures, the critical matrix verbs and their infinitival complements appeared in subordinate clauses. Examples (7) and (8) illustrate the critical experimental conditions for both acceptability judgement experiments. Each experimental item set included two further word order variants, which however are irrelevant to the current study and thus will not be analysed or discussed.

- (7) Example test item set for Experiment 2A with the verb *versuchen* ('try')
- a. EXTRAPOSITION  
*Julia sagt, dass Fred versucht [ihn zu streicheln].*  
Julia says that Fred tries him to pet
  - b. THIRD CONSTRUCTION  
*Julia sagt, dass Fred [ihn] versucht [zu streicheln].*  
Julia says that Fred him tries to pet  
'Julia says that Fred tries to pet him.'
- (8) Example test item set for Experiment 2B with the verb *versuchen* ('try')
- a. EXTRAPOSITION  
*Finn sagt, dass der Junge versucht [sich die Hände zu waschen].*  
Finn says that the boy tries himself the hands to wash

b. INTRAPOSITION AMBIGUOUS

*Finn sagt, dass der Junge [sich die Hände zu waschen versucht].*

Finn says that the boy himself the hands to wash tries

‘Finn says that the boy tries to wash his hands.’

In the Extraposition conditions (7a) and (8a), the infinitival complement follows its subcategorising verb and is thus placed at the right periphery of the sentence. In condition (8b) the infinitival complement is intraposed between the subject noun phrase *Fred* and the matrix verb *versucht*. Finally, the Third Construction (7b) represents a discontinuous construction, in which the accusative-marked object pronoun *ihn* appears before the matrix verb whilst the verbal phrase *zu streicheln* appears in an extraposed position to its right. For both experiments we created 16 critical item sets by placing each of the four matrix verbs in four different contexts. Since the extraposed condition was tested in both Experiment 2A and 2B, it actually occurred in eight different contexts. All experimental sentences began with an introductory clause such as *Julia sagt,...* (‘Julia says...’) followed by the conjunction *dass* (‘that’) introducing a subordinate clause.

In addition, both experiments included filler sentences (Experiment 2A: 40 filler sentences; Experiment 2B: 16 filler sentences) which did not include any infinitival complements and differed from the experimental items both in their lexical material and syntactic structure. On the one hand, fully acceptable fillers such as *Klaus glaubt, dass die Erklärungen des Lehrers nicht richtig sind*. (‘Klaus believes that the teacher's explanations are incorrect.’) were included. On the other hand, ungrammatical ‘sanity-check’ filler sentences such as *\*Lena prüft, was verursacht beim Unfall für Schäden sie hat*. (lit. ‘Lena is checking what caused in the accident for damage she has.’) allowed us to verify whether participants read the stimulus sentences attentively, followed the experimental task and made use of the full range of the judgement scale. In Experiment 2A, a picture illustrating the entity referred to by the personal pronoun in the subordinate clause (e.g. a picture of a dog in (7a) and (7b)) was presented above the stimulus sentence so as to provide a plausible referent for the pronoun and thus to make the sentences more easily interpretable. In Experiment 2B, our use of the reflexive pronoun *sich* made it unnecessary to present a context picture. For both experiments, four experimental lists were created using a Latin square design, so that each participant saw each item in each context only once. Critical and filler items were presented in pseudo-randomized order with no more than three critical items presented in a row.

**Procedure**

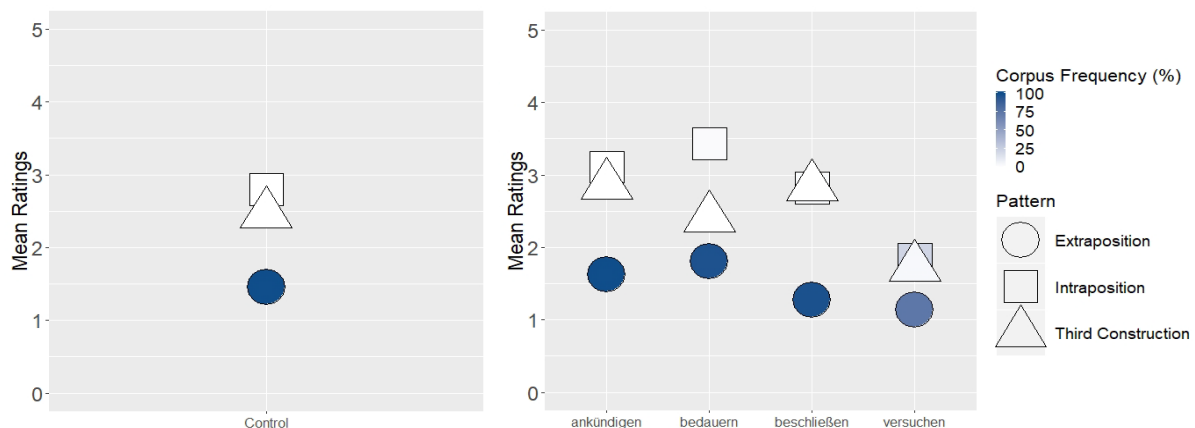
Both experiments were administered via the world-wide web, such that participants received a link to a web-based questionnaire implemented via Google forms (Google Inc., USA). The questionnaire each consisted of three parts: Firstly, participants were asked to answer a few biographical questions as well as to give their consent to the participation in the study. Secondly, following the instructions, three practice items were presented to allow participants to familiarize themselves with the experimental task. After giving participants the chance to ask any remaining questions about the experimental procedure and task, the experiment started with the presentation of the first trial.

Participants were asked to read each stimulus sentence carefully and then rate their acceptability on a 5-point Likert scale. Sentences were presented individually in Arial 20pt in black colour against a white background, with the Likert scale appearing right underneath the test sentence. The Likert scale ranged from 1 labelled as “vollkommen akzeptabel” (‘totally acceptable’) to 5 labelled as “völlig inakzeptabel” (‘totally unacceptable’), resembling the grading scale commonly used in German schools and universities. Points 2, 3, and 4 were not labelled on every single scale, but were introduced as 2 meaning “eher akzeptabel” (‘rather

acceptable’), as 3 meaning “nicht sehr akzeptabel” (‘not very acceptable’), and as 4 meaning “eher inakzeptabel” (‘rather unacceptable’). For each sentence, participants could take as much time as they needed. Once they had provided their acceptability judgement for a given sentence, the next trial started automatically. Experiment 2A took approximately 25 minutes to complete and the somewhat shorter questionnaire in Experiment 2B took 15 minutes to complete on average.

## Results

Acceptability ratings were z-transformed and linear mixed-effects models were conducted using R (Baayen et al., 2008; R Core Team, 2017). In order to perform direct comparisons of rating patterns across all three word order variants under investigation, i.e. Extraposition (7a) and (8a), Intraposition (8b), and Third Construction (7b), data from both Experiments 2A and 2B were combined for statistical analyses. Since the extraposed condition was tested in both Experiment 2A and 2B, it produced twice as many data points as the other two experimental conditions. The factorial structure of our experiment was reflected in the structure of our models, such that the factors ‘Condition’ (comparing the three word order variants under investigation), and ‘Verb’ (examining effects of word order variation across individual matrix verbs) as well as interactions between those factors were included. Based on the log-likelihood comparison, the model that best fit the data, starting with the maximal random-effects structure and removing them successively, was determined at the .05 level. Contrast coding was employed for the factorial predictors and the comparisons of interest were obtained by releveling factors and refitting the model.



**Figure 2.** Mean acceptability ratings for the three word order patterns across all control verbs (left panel) and for individual verbs (right panel), Experiment 2

Figure 2 shows mean acceptability ratings for the different word order variants across all control verbs (left panel) and for individual control verbs (right panel) in relation to the relative frequency of each word order variant as indicated by colour shading, with dark shading indicating ‘high frequency’ and white indicating ‘unattested’. Across all control verbs, extraposed infinitival complements were preferred over intraposed infinitival complements and Third Constructions, while Intraposition and Third Construction produced similar acceptability ratings. Statistical analyses confirmed this main effect of condition. The most frequent word order variant, namely Extraposition, was indeed rated most acceptable for all control verbs under investigation, while ratings for Intraposition and Third Construction did not statistically differ (*Extraposition vs. Intraposition*:  $b = 0.8589$ ;  $t = 12.318$ ;  $p < .001$ ; *Extraposition vs. Third Construction*:  $b = 0.7946$ ;  $t = 8.399$ ;  $p < .001$ ; *Intraposition vs. Third Construction*:  $b = -0.0643$ ;  $t = -0.550$ ;  $p = .582$ ).

In addition, statistical analyses revealed significant interactions between the factors Condition and Verb, indicating that speakers' word order preferences were modulated by the individual control verb they encountered. Word order patterns which are infrequent or even unattested, i.e. Intraposition and Third Construction, exhibit substantial variability in rating patterns and were generally rated much more favourably than their frequency counts would suggest. In particular, intraposed infinitival complements as well as those in the Third Construction for the control verbs *ankündigen* and *beschließen* elicited unexpectedly high acceptability ratings in the intermediate scalar range, although they were highly infrequent or even unattested in both our corpus data and the spoken production data. Additionally, on the one hand, patterns which are equally infrequent exhibit significant variability in their ratings, such as Intraposition and Third Construction for the control verb *bedauern* ( $b = -0.5248$ ;  $t = -2.237$ ;  $p = .026$ ). On the other hand, patterns which show significantly different frequency distributions, such as Intraposition and Third Construction for the control verb *versuchen*, yield very similar acceptability ratings ( $b = 0.0161$ ;  $t = 0.102$ ;  $p = .918$ ). These findings suggest that the frequency distributions of the different word order variants for the four infinitive-embedding matrix verbs under scrutiny do not fully correlate with speakers' acceptability judgement patterns.

### 3.3 Discussion: Frequency vs. Language Performance

Testing the predictions made by usage-based accounts to word order variation in language performance and representation revealed a manifold picture: On the one hand, we found that in our spoken production task, speakers' word order preferences indeed correlate with corpus probabilities. One potential explanation for the consistent findings from our corpus analysis and the spoken production experiment might be rooted in the types of datasets involved. Frequency distributions as obtained from corpus data essentially reflect native speakers' production choices. Similarly, in elicited spoken production the speaker also evaluates a set of alternative, but potentially equally suitable, word order variants and ultimately chooses the preferred candidate as part of a forced selection process. Therefore, corpus distributions and production preferences can be considered similar types of data.

On the other hand, however, we observed discrepancies between corpus frequencies / spoken productions and the findings from our acceptability judgement experiments. These discrepancies suggest a two-fold answer to the question of whether and to what extent statistical frequency correlates with speaker performance: Only for high-frequency patterns does performance match corpus probabilities, while for low-frequency patterns, performance cannot solely be accounted for by frequency distributions. That is, we obtained both a *ceiling effect*, such that only the most frequent word order variant, namely Extraposition, also received the best ratings across all four control verbs, and a *floor mismatch effect* (Bader & Häussler, 2010), such that low frequency word order variants, i.e. Intraposition and the Third Construction, yielded substantial variability in rating patterns relative to their frequency counts. According to Bader & Häussler (2010), a floor mismatch effect can be observed if two or more syntactic variants differ in terms of perceived well-formedness although their frequencies of occurrence are approaching zero level (p. 316).

The observation that infrequent and/or unattested syntactic variants receive fairly favourable ratings in judgement tasks has been discussed both from the perspective of linguistic theory-building and from a psycholinguistic perspective. Newmeyer (2003), for example, claimed that language users are able to judge syntactic constructions as acceptable alternatives even though they would never make use of them in language production, probably because these structures are present as latent structures in the speakers' grammar. In line with Newmeyer's (2003) observations, Culicover (2013) also proposes that syntactic variants that

are hardly encountered in language use but are judged as acceptable “may reside in a speaker without overt symptoms [such as viruses], that is, they are accepted as ‘possible’ or ‘acceptable’, but they are not produced” (p.252). Accordingly, any asymmetry between the outcome of repeatedly produced forced-choice selections, i.e. corpus or elicited production data, and speakers’ acceptability judgements of syntactic variants can be attributed to a power-law distribution, such that most acceptable variants are produced consistently, while less acceptable but still potentially felicitous candidates are hardly produced at all (Featherston, 2005; Verhoeven & Temme, 2017).

### 3.4 Experiment 3: Sentence Reading Times

Since comparing our production and judgement data revealed that frequency distributions of different word order variants in German infinitival complementation do not fully correlate with speakers’ performance patterns, particularly for low-frequency alternatives, the question remains whether and to what extent processing-related factors might affect word order preferences. Hence, we conducted a timed reading comprehension task (e.g. Frazier et al., 1983) to measure the relative processing cost of the three different word order variants: Extraposition, Intraposition and Third Construction. Increased processing difficulty for particular structural variants is assumed to be reflected in elevated reading times relative to easier-to-process word order patterns. Based on processing-based constraints which attempt to capture the sources of processing cost in syntactic variation contexts (Hawkins, 1994; Gibson, 2000), we expect centre-embedded and discontinuous construals, i.e. Intraposition and the Third Construction, to elicit elevated reading times compared to non-nested extraposed infinitival complements across all infinitive-embedding control verbs under investigation.

#### *Participants*

The reading-time study tested 48 adult native speakers of German (11 male, 37 female) aged between 18 and 37 years (mean: 25.21 years). Similarly to the other experimental studies, all participants were recruited from among the student and working communities in and around Potsdam and/or Berlin and reported not to have any background in linguistics. None of the participants reported to suffer from any language and/or reading disorders. They were not informed about the ultimate purpose of the study before testing and received 4€ as reimbursement for their participation.

#### *Materials*

Experimental materials were created by combining each of the four chosen matrix verbs with corresponding infinitival complements in the three different word order variants. To make sure that infinitival complements preserve SOV structure, they were again presented in subordinate clauses. 16 critical items sets were created by using each of the four chosen matrix verbs in four different sentence contexts. The experimental item set shown in (9) exemplifies one such context.

- (9) Example test item with the verb *versuchen* (‘try’)
- a. EXTRAPOSITION  
*Julia sagt, dass Fred versucht [den Kuchen zu schneiden].*  
Julia says that Fred tries the cake to cut
  - b. INTRAPOSITION  
*Julia sagt, dass Fred [den Kuchen zu schneiden] versucht.*  
Julia says that Fred the cake to cut tries
  - c. THIRD CONSTRUCTION  
*Julia sagt, dass Fred [den Kuchen] versucht [zu schneiden].*  
Julia says that Fred the cake tries to cut

‘Julia says that Fred tries to cut the cake.’

Each experimental sentence began with an introductory clause such as *Julia sagt, ...* ('Julia says...'), followed by the conjunction *dass* ('that') and the rest of the critical subordinate clause. Condition (9a) contains an extraposed infinitival complement, a structural variant which leaves the subject-verb dependency between *Fred* and *versucht* uninterrupted. In experimental condition (9b), the infinitival complement is intraposed between the subject *Fred* and the matrix predicate *versucht*, thus increasing subject-verb distance by means of centre-embedding. Finally, the Third Construction in (9c) represents a discontinuous construction in which the accusative complement of the verb *schneiden* 'cut' (*den Kuchen* 'the cake') appears in front of the matrix verb *versucht*, while the remainder of the infinitival complement is extraposed to the right of the matrix verb. This condition does not only break up two dependencies, i.e. the matrix subject-verb dependency and the infinitival complement itself, but it is also locally ambiguous as the accusative-marked preverbal noun phrase might be analysed as a direct object of the matrix verb initially. Hence processing Third Constructions successfully may require some structural reanalysis. Revising an initial misanalysis might be particularly effortful for plausible relative to implausible noun phrases. Therefore, the present experimental design controlled for plausibility by ensuring that for each of the four verbs under investigation, two of the four experimental sentences included plausible object noun phrases (cf. example (9)), while the two remaining test sentences included noun phrases that were implausible as direct objects of the matrix verb (e.g. *...dass Hans den Traktor versucht zu bremsen* ‘...that Hans tries to slow down the tractor’).

In addition, 32 filler sentences, such as “*Klaus glaubt, dass die Erklärungen des Lehrers nicht richtig sind.*” ('Klaus believes that the explanations of the teacher are incorrect.'), were created in order to hide the real purpose of the experiment. Filler sentences differed from experimental sentences in their lexical material and in their syntactic structure, such that infinitival complements were avoided completely. In addition, after one-third of our stimulus items, participants were asked to answer corresponding comprehension questions (example question for (9): *Versucht Fred das Brot zu schneiden?* ‘Does Fred try to cut the bread?’) which allowed us to make sure that participants attentively read our stimulus sentences and followed the experimental task. Three experimental lists were created making use of a Latin Square Design. Critical and filler items were presented in pseudo-randomized order with no more than three critical items presented in a row.

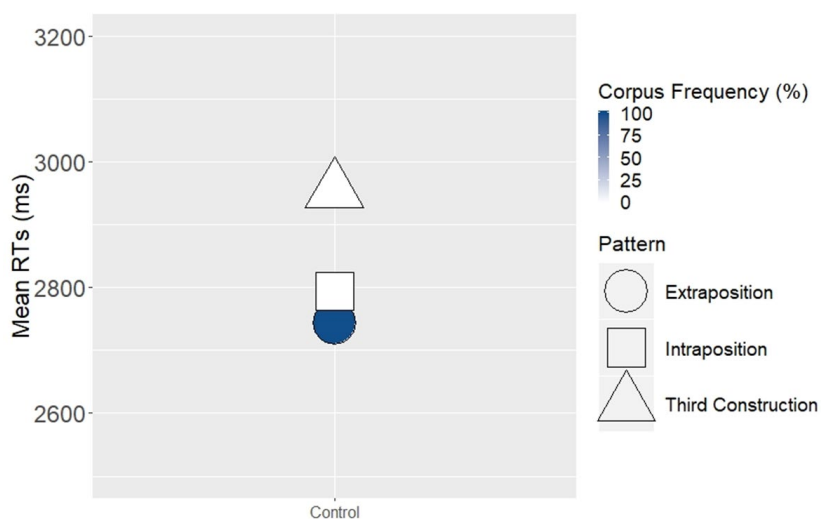
### **Procedure**

The reading-time task was administered via the internet. The experiment was implemented on the web-based experimental platform Ibx Farm (Drummond, 2013) and participants received a link to it. After answering a set of biographical questions, the experiment itself started with the presentation of three practice items. The presentation of each item began with a fixation line and as soon as participants pressed the spacebar, the test item appeared on the computer screen. Each item was presented individually as a full sentence in Arial 20pt in black colour against a light grey background. Participants were asked to read each stimulus sentence as quickly as possible for meaning and to indicate as soon as they had finished reading each sentence by clicking on the spacebar. Participants were instructed to attend carefully to the sentences' meaning, since after one-third of the stimulus items a *yes/no* comprehension question appeared on the screen and participants had to indicate their answer by clicking on a ‘YES’ or ‘NO’ button. Both full-sentence reading times and accuracy scores for comprehension questions were recorded. After answering the question, the presentation of the next trial started automatically by showing the fixation line for the next sentence. On average, the reading-time experiment took approximately 30 minutes to complete.

## Results

Overall, accuracy scores were close to ceiling with 95.31 % correct answers (by-participant range: 75 %-100 %; by-item range: 80 %-100 %), thus no participant or item had to be excluded from further statistical analyses. The reading-time data was log-transformed for statistical analyses (Baayen & Milin, 2010) and linear-mixed effects models were conducted using R (Baayen et al., 2008; R Core Team, 2017). The factorial structure of the experiment was reflected in the structure of our models, including the factors ‘Condition’ and ‘Verb’ as well as interactions between them. Based on the log-likelihood comparison, the model that best fit the data, starting with the maximal random-effects structure and removing them successively, was determined at the .05 level. Contrast coding was employed for the factorial predictors and the comparisons of interest were obtained by releveling factors and refitting the model.

Figure 3 shows mean reading times for the different word order variants across all control verbs in relation to the relative frequency of each word order variant as indicated by colour shading, with dark shading indicating ‘high frequency’ and white indicating ‘unattested’. Across all control verbs, extraposed and intraposed infinitival complements elicited shorter reading times than Third Constructions, with Intraposition and Extraposition patterns eliciting similar reading times. Accordingly, statistical analyses revealed a significant effect of the factor ‘Condition’, such that the Third Construction elicited significantly longer mean reading times than extraposed and intraposed infinitival complements (*Third Construction vs. Extraposition*:  $b = -0.0706$ ;  $t = -2.152$ ;  $p = .032$ ; *Third Construction vs. Intraposition*:  $b = -0.0673$ ;  $t = -2.046$ ;  $p = .042$ ), while mean reading times for the latter two did not statistically differ (*Extraposition vs. Intraposition*:  $b = 0.0033$ ;  $t = 0.102$ ;  $p = .918$ ). However, differences between the three word order patterns were not significantly modulated by individual control verbs (all  $p > .05$ ).



**Figure 3.** Mean reading times for the three word order patterns across control verbs, Experiment 3

Thus, infinitival complements in the Third Construction were found to incur significantly more processing cost during sentence comprehension compared to the other two word order variants.

### 3.5 Discussion: Processing Economy vs. Language Performance

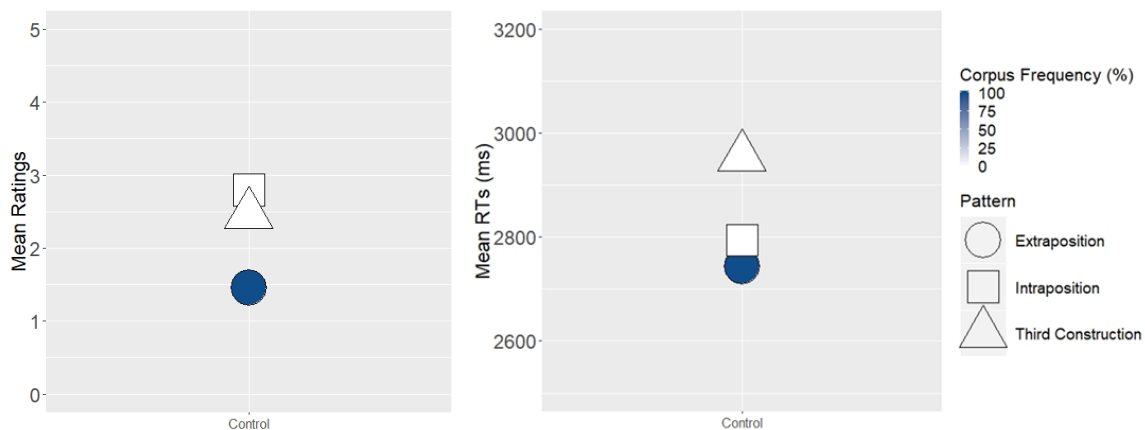
The results from Experiment 3 revealed different response patterns compared to the other experimental tasks. Firstly, extraposed infinitival complements, together with intraposed structures, elicited the fastest reading times. Secondly, we saw a significant processing disadvantage for Third Constructions relative to the other two word order patterns. The finding that extraposed infinitival complements elicited the fastest reading times is in line with previous proposals of processing economy. From the perspective of both the principle of *Closure*



(Prideau & Baker, 1986), which favours word order patterns that allow clausal units to be completed as soon as possible during language comprehension, and the principle of *Early Immediate Constituents* (Hawkins, 1994), which gives preference to word order patterns that allow a rapid and efficient recognition of syntactic groupings, a processing advantage for extraposed infinitival complements was expected.

The observed reading time pattern also confirms the predictions made by Gibson's *Dependency Locality Theory* (2000), which suggests to quantify processing difficulty in relation to the distance between structurally dependent elements. In this respect, the Third Construction is the only word order pattern that breaks up two structural dependencies, namely the matrix subject-verb dependency and the infinitival complement itself (cf. example (9c)). In addition, Third Constructions are locally ambiguous in that the accusative noun phrase preceding the matrix predicate might initially be analysed as its direct object, creating a need for the initial interpretation to be revised upon encountering the infinitival phrase. As has been shown for other types of locally ambiguous ('garden-path') sentences, this reanalysis may be associated with additional processing effort relative to structurally simpler constructions, such as intraposed and extraposed infinitives. Thus, a measurable increase in reading times was expected for Third Constructions.

Furthermore, the significant difference in reading times for intraposed infinitival complements and Third Constructions yields an interesting asymmetry of findings for these two low-frequency structures: Although both word order variants are similarly infrequent according to both corpus frequencies and elicited productions and yielded similar judgement patterns in the intermediate scalar range in both of our ratings tasks, full-sentence reading times suggest substantially increased processing difficulty for the Third Construction relative to Intraposition (see Figure 4).



**Figure 4.** Mean acceptability judgements (left panel) versus mean reading times (right panel) for the three word order patterns (averaged across control verbs)

The significantly higher processing cost incurred by Third Constructions vs. the two other construction types under investigation might help explain why Third Constructions are highly dispreferred in actual language use. Our finding that language users are nevertheless able to provide reliable - and not particularly unfavourable - judgements for these structures suggests that the Third Construction is nevertheless present in German speakers' grammar as a latent structure, however (compare Newmeyer, 2003; Culicover, 2013). What remains puzzling about the Third Construction is that, although it is associated with highest processing costs, this construction is attested almost exclusively in spoken or spoken oriented (e.g. Twitter) language. One possible explanation could be that the Third Construction signals a particular information structure, and that information-structural needs override processing difficulties in spoken

language. According to Cook (2001), discontinuous infinitival complements are used when the non-finite verb and its complement have different information status. Considering the tendency of topical information to appear as high as possible in the sentence, one could hypothesize that topical parts of the infinitival complement are realised pre-verbally, while other, possibly new parts are realised post-verbally. A first glance at our data suggests that this might indeed be the case.

Intraposited (i.e. centre-embedded) infinitival complements were also hardly ever used in language production, but unexpectedly they did not elicit significantly slower full-sentence reading times compared to extraposed structures. This suggests that Intraposition did not incur any additional processing cost relative to non-nested construals. This might be at least partly due to Intraposition potentially allowing for a monoclausal construal. Bayer et al. (2005) report experimental evidence indicating that sentences containing intraposed infinitives are preferentially construed as monoclausal, which would reduce overall sentence complexity relative to a biclausal construal and thus likely reduce processing cost. Note, however, that Experiment 3 only measured full-sentence reading times in order to capture differences in comprehension ease between the word order patterns under investigation. It could also be the case that due to their relative shortness, the centre-embedded structures we presented in the intraposed condition were still comparatively easy to process and did not present a sufficient contrast in terms of structural complexity relative to their extraposed counterparts. Consequently, it is possible that full-sentence reading time measures were not able to capture potential processing differences during real-time comprehension between these two word order variants. Clearly, additional and more fine-grained research on the processing effort involved in the computation of these structures by means of more time-sensitive measures is necessary.

## **4 Conclusion**

To shed more light on the constraints which limit word order variation in German infinitival complementation, the present study investigated two frequently discussed factors: (i) frequency of occurrence and (ii) processing ease. Firstly, our multi-methodological approach to syntactic variation has shown that frequency-based approaches to language processing and representation do not fully account for speakers' preference patterns, as frequency counts only capture the preference of high-frequency structures but fail to account for variability in performance for low-frequency construals. Secondly, the results from the present study indicate that processing cost, especially for the Third Construction, can influence speakers' preference patterns across a variety of tasks in language production and judgement. Consequently, the present study presents a case for taking into account experimental data from different sources to obtain a more comprehensive and nuanced picture of grammatical preferences in variation contexts.

## **Acknowledgments**

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317633480 – SFB 1287, Project A02. We particularly thank Lara Teich for creating the picture card stimuli used in Experiment 1, our graduate student, Patricia Weber, for the provision of the data in Experiment 2B as well as our student assistant, Birte Pravemann, for her great support in collecting and analysing both corpus and behavioural data.

## References

- Baayen, H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research* 3(2), 12–28.
- Baayen, H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.
- Bader, M., & Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics* 46(2), 273–330.
- Bayer, J., Schmid, T., & Bader, M. (2005). Clause union and clausal position. In M. den Dikken & C. Tortora (Eds.), *The Function of Function Words and Functional Categories* (pp. 79–113). Amsterdam: John Benjamins.
- Bech, G. (1983[1955]). *Studien über das deutsche verbum infinitum*. Tübingen: Niemeyer.
- Bosch, S., De Cesare, I., Demske, U., & Felser, C. (in press). Word order variation and coherence in German infinitival complementation. *The Journal of Comparative Germanic Linguistics*.
- Brandner, E. (2020). A “borderline case” of syntactic variation. *Glossa: a journal of general linguistics*, 5(1), 25. DOI: <http://doi.org/10.5334/gjgl.606>
- Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language* 82(4), 711–733.
- Bybee, J., & Beckner, C. (2010). Usage-based theory. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 827–856). Oxford: Oxford University Press.
- Bybee, J., & Hopper, P. (2001). Introduction to frequency and the emergence of linguistic structure. *Typological Studies in Language* 45, 1–26.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry* 36(1), 1–22.
- Cook, P. (2001). Coherence in German: An information structure approach. Doctoral Dissertation, University of Manchester.
- Culicover, P. (2013). *Explaining Syntax. Representations, Structures, and Computation*. Oxford: Oxford University Press.
- Drummond, A. (2013). Ibox Farm. <http://spellout.net/iboxfarm/>.
- Featherston, S. (2005). The decathlon model of empirical syntax. In M. Reis & S. Kepsner (Eds.), *Linguistic Evidence* (pp. 187–208). Berlin: Mouton de Gruyter.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14(2), 178–210.
- Frazier, L., Clifton, C., & Randall, J. (1983). Filling gaps: Decision principles and structure in sentence comprehension. *Cognition*, 13(2), 187–222.
- Gahl, S., & Garnsey, S. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Linguistic Society of America* 80(4), 748–775.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1), 1–76.
- Gibson, E. (2000). The dependency-locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. P. Marantz & W. O'Neill (Eds.), *Image, Language, Brain* (pp. 95–126). Cambridge, MA: MIT Press.
- Grano, T. (2011). Mental action and event structure in the semantics of 'try'. In *Proceedings of Semantics and Linguistic Theory 21* (pp. 426–443). DOI: <https://doi.org/10.3765/salt.v21i0.2607>
- Haider, H. (2010). *The Syntax of German*. Cambridge: Cambridge University Press.
- Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency* (Vol. 73). Cambridge: Cambridge University Press.

- Kempen, G., & Harbusch, K. (2005). The relationship between grammaticality ratings and corpus frequencies. A case study into word order variability in the middlefield of German clauses. In M. Reis & S. Kepser (Eds.), *Linguistic Evidence – Empirical, Theoretical, and Computational Perspectives* (pp. 329–349). Berlin: Mouton De Gruyter.
- Larsen-Freeman, D. (1997). Chaos/Complexity science and second language acquisition. *Applied Linguistics* 18, 141–165.
- Newmeyer, F. (2003). Grammar is grammar and usage is usage. *Language* 79(4), 682–707.
- Prideaux, G. D., & Baker, W. J. (1986). *Strategies and Structures: The Processing of Relative Clauses*, Amsterdam, Philadelphia: John Benjamins.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Radford, A., Felser, C., & Boxell, O. (2012). Preposition copying and pruning in present-day English. *English Language and Linguistics* 16, 403–426.
- Reis, M. (2001). Bilden Modalverben im Deutschen eine syntaktische Klasse? In R. Müller & M. Reis (Eds.), *Modalität und Modalverben im Deutschen. Linguistische Berichte* 9 (pp. 287–318). Hamburg: Helmut Buske Verlag.
- Ritter, E., & Wiltschko, M. (2014). The composition of INFL. *Natural Language & Linguistic Theory*, 32(4), 1331–1386.
- Schmid, T., Bader, M. & Bayer, J. (2005). Coherence - An experimental approach. In M. Reis & S. Kepser (Eds.), *Linguistic Evidence. Empirical, Theoretical and Computational Perspectives* (pp. 435–456). Berlin: Mouton De Gruyter.
- Verhoeven, E., & Temme, A. (2017). Word order acceptability and word order choice. In S. Featherston, R. Hörnig, R. Steinberg, B. Umbreit & J. Wallis (Eds.), *Proceedings of Linguistic Evidence 2016: Empirical, Theoretical, and Computational Perspectives* (pp. 1–17). Tübingen: University of Tübingen. <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77066>.
- Wurmbrand, S. (2001). *Infinitives. Restructuring and Clause Structure*. Berlin: Mouton de Gruyter.
- Wurmbrand, S. (2014). Restructuring across the world. In L. Veselovská & M. Janebová (Eds.), *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*, (pp. 275–294). Olomouc: Palacký University.