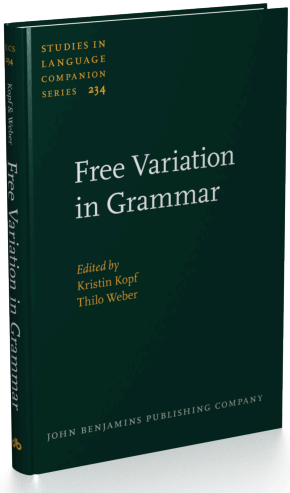
Studies in Language Companion Series
234

CHAPTER 2

How free is the position of German object pronouns?

Markus Bader | Goethe University Frankfurt

https://doi.org/10.1075/slcs.234.02bad



Pages 22–47 of
Free Variation in Grammar: Empirical and theoretical approaches
Edited by Kristin Kopf and Thilo Weber
[Studies in Language Companion Series, 234]
2023. vi, 352 pp.

CHAPTER 2

How free is the position of German object pronouns?

Markus Bader
Goethe University Frankfurt

Corpus studies show that weak object pronouns in German directly precede or follow the subject, depending on properties of the subject, including weight, animacy and thematic role. Whether the same factors also affect the acceptability of sentences with object pronouns was investigated in three magnitude-estimation experiments. The results show that both orders (object pronoun before/after subject) are highly acceptable, with some small acceptability differences depending on weight, animacy and thematic roles. Based on these results, the hypothesis is advanced that the position of weak object pronouns in German relative to the subject is an instance of free variation within the grammar but choosing a specific order during language production follows general production preferences and is thus not random.

Keywords: word order, object pronouns, German, magnitude estimation, language production

1. Introduction

From a syntactic point of view, speakers of German enjoy much freedom when it comes to the order of subject and object(s). As soon as non-syntactic factors are taken into account, however, the apparent word-order freedom often disappears. Such non-syntactic factors derive from, among others, conceptual structure, information structure, and the weight of phrasal constituents. For example, when a sentence contains an agentive verb, an animate proper name as subject, and an inanimate indefinite NP as object, it is hardly possible to put the object in front of the subject within the so-called middlefield (e.g., ?? *Vermutlich pflückte einen Apfel Peter*. 'Presumably, Peter picked an apple').[1] Constraints on word

1. According to the topological model of German sentence structure (Drach 1937), the *middlefield* is that part of a sentence that starts directly after the finite verb in a verb-second clause and

https://doi.org/10.1075/slcs.234.02bad

order freedom have been amply studied within theoretical, experimental, and corpus linguistics (see overview in Weskott 2021). This research has shown that, in many cases, one order can be used quite freely (typically, but not always, subject-before-object order), whereas the use of alternative orders is severely restricted.

There is one area of German syntax, however, where subject-before-object (SO) and object-before-subject (OS) order seem to be truly exchangeable: pronominal objects can freely occur before or after a non-pronominal subject. Compare the corpus examples in (1a) and (1b).

(1) a. Der Mann stürzte dadurch zu Boden, so dass *der Detektiv* **ihn**
the man fell thereby to ground so that the detective him
überwältigen und bis zum Eintreffen der Polizei festhalten konnte.
overpower and until the arrival the police detain could
'Thereby, the man fell to the ground so that the detective could overpower him and detain him until the police arrived.'
https://www.rhempfalz.de/pfalz_artikel,-kaiserslautern-ladendieb-z%C3%BCckt-messer-_arid,337877.html

b. Beim Sprung über einen Maschendrahtzaun verletzte sich der Mann,
at jump over a chain-wire fence hurt himself the man
so dass **ihn** *der Detektiv* einholen und festhalten konnte.
so that him the detective catch and detain could
'When jumping over a chain-wire fence, the man hurt himself, so that the detective could catch and detain him.'
https://www.come-on.de/luedenscheid/mann-jagt-raeuber-durchs-stern-center-luedenscheid-wegen-schutzmasken-13809523.html

The two examples in (1a) and (1b) are quite similar with regard to properties that are known to affect word order: in both sentences, the embedded clause contains an animate subject, an animate object, and an agentive verb. Furthermore, the antecedent of the object pronoun is contained in the preceding main clause, and the subject is a definite NP that refers to a detective mentioned in the preceding context (not shown here). Despite this similarity, we find SO order in the embedded clause of (1a) but OS order in the embedded clause of (1b). In both cases, switching to the alternative order seems to be possible without any change in meaning and any loss of acceptability. The impression that the order of subject and object is not subject to grammatical constraints when the object is a pronoun is reinforced by a look into the syntactic literature on word order variation in German. For example, Müller (1999) and Haider (2010, Chapter 4) provide sophisticated discussions of the conditions under which an object can precede the subject

after the complementiser in a verb-final clause, and ends directly before the verb(s) in clause-final position.

in the middlefield. Object pronouns, however, are exempted from the discussion, under the premise that the serialisation of subject and object pronoun is a matter of free variation.

As the editors make clear in their introduction to this volume, free variation can be understood in different ways. The discussion so far has considered free variation within the grammar, that is, sentences that can be realised with different word orders without any effect on the meaning of the sentence – with 'meaning' being understood in a very broad sense, including semantic, pragmatic and social aspects. Even if two word order variants are free in grammatical terms, they are not necessarily free in an absolute sense, that is, the possible orders – SO and OS in the case under consideration – are not necessarily chosen randomly during language production. As shown by corpus studies that will be reviewed in the next section, the position of object pronouns in the German middlefield is clearly not free in this absolute sense. Instead, the likelihood that one of the two orders is produced depends on a range of word order preferences that have been firmly established in typological research, including such well-known principles as 'animate before inanimate' or 'short before long'. The two examples in (1a) and (1b) come out as about equally probable when subjected to the preference rules that are revealed by corpus studies. For example, an agentive subject favours SO order, whereas a definite subject favours OS order. When a sentence contains a subject that is agentive and definite, the likelihood of producing a sentence with SO order can, therefore, be about the same as the likelihood of producing a sentence with OS order, as for example in (1a) and (1b).

In other cases, constraints jointly pull in a single direction, resulting in a strong preference for SO or OS order. For example, when the object pronoun is animate (which it was in the overwhelming majority of sentences in the corpus study of Bader 2020; see below) but the subject inanimate, a strong preference for OS order is observed, mirroring the well-known preference for 'animate-before-inanimate' orders found with non-pronominal NPs in German (see Hoberg 1981; Bader and Häussler 2010b). However, even in such cases, authentic examples for both of the two possible orders can be found, as illustrated in (2).

(2) a. Auf jeden Fall ist dem Biologen anzumerken, dass **ihn** *seine neue Arbeit* fasziniert.
'In any case, the biologist shows that his new job fascinates him.'
https://www.saechsische.de/vom-labor-ins-klassenzimmer-3485259.html

b. Im Interview gab ROLAND zu Protokoll, dass *die Rolle* **ihn** fasziniert habe
'In the interview, ROLAND put on record that the role fascinated him.'
https://schlagerprofis.de/32910/

The question that is addressed in this paper is whether SO and OS order are generally of equal, high acceptability in the case of object pronouns, or whether this only holds when both orders occur with about equal frequency in language production but not when one order strongly outnumbers the other. If no acceptability differences show up, we will be entitled to conclude that the order between pronominal object and non-pronominal subject is truly free from a grammatical perspective, with all frequency differences that have been found being a matter of language production. If, on the other hand, acceptability differences between the two orders should be found under certain conditions, drawing conclusions will be more involved. As has been discussed by many authors, differences in acceptability can reflect differences in grammaticality but they can also be caused by the processing mechanisms, stemming, for example, from differences in processing complexity (see Fanselow 2021, for a recent overview). Since it is not possible to decide a priori between grammar- and processing-based explanations of acceptability differences, a more thorough discussion of this issue must wait until the results of the upcoming experiments have been presented.[2]

The organisation of this chapter is as follows. Section 2 reviews corpus studies of the position of object pronouns in German. This allows us to identify configurations in which OS or SO order is strongly preferred to the alternative order. Based on the corpus results reviewed in Section 2, three acceptability experiments have tested whether sentences with a pronominal object differ in acceptability depending on whether the sentences occur with SO or OS order. These experiments are presented in Section 3. The paper concludes with an evaluation of the experimental results in the light of known corpus preferences in Section 4.

2. What governs the position of object pronouns?

The position of object pronouns in the German middlefield has been the subject of several corpus studies leading to converging results (Hoberg 1981; Heylen 2005; Kempen and Harbusch 2005; Bader 2020). The following summary is based on Bader (2020) because that study analysed the largest set of examples, namely, a set of 4,322 sentences containing object pronouns. All sentences were randomly drawn from the deWac corpus (Baroni et al. 2009), a large collection of internet texts of all sorts.

Based on the complete set of 4,322 examples, Bader (2020) extended the word order template of the German middlefield proposed in Hoberg (1981), resulting

2. Prescriptive biases are a further important source of acceptability variation, as shown by Vogel (2019). As far as I can see, this is of no relevance to the sentences at hand.

in the word order template shown in (3). As indicated, this template specifies the position of weak object pronouns. Object pronouns that are strong because they are focused and, therefore, accented are not constrained in the same way and can appear where other NPs can appear too. In the following, the term 'object pronoun' will always be understood as referring to weak object pronouns.

(3) Basic word order template for the German middlefield
(N/A/D = NP with nominative/accusative/dative case)

(N)	(A-D)	((N-A-D)+ani-(N-A-D)-ani)	(N, D, A)
Subject	Weak object pronouns	Non-pronominal NPs, adverbials	Idioms

According to this template, weak object pronouns immediately follow the subject when it is in the initial position of the middlefield. Because the subject can also appear in the region following the position of object pronouns, the position of object pronouns relative to the subject is variable. Object pronouns appear in the initial position of the middlefield when the subject occurs at any later position, or they directly follow the subject when it occupies the initial position of the middlefield. Later positions, in contrast, are excluded for object pronouns according to this template. The position of object pronouns within the middlefield is thus not completely free – object pronouns must occur early, but how early is a matter of variation.

This variation is not random but follows – in a probabilistic way – the major linguistic hierarchies identified in typological research to govern the choice between alternative word orders (see overviews in Siewierska 1993; Croft 2003). For the German middlefield, of particular relevance is the animacy hierarchy (animate < inanimate), but other hierarchies are also at work, including the definiteness hierarchy (definite < indefinite), the givenness hierarchy (given < new), the thematic role hierarchy (agent < non-agent) and the case hierarchy (nominative < non-nominative) (see overviews in Hoberg 1981; Lerot 1985; Müller 1999; Bader and Häussier 2010b).

For the following experiments, two hierarchies from the hierarchies investigated in the corpus study of Bader (2020) were selected for investigation. All percentages cited below are from this corpus study. The first hierarchy is the animacy hierarchy as applied to the subject NP. Overall, sentences with an animate subject occurred with OS order in 49% of all cases, whereas 85% OS order was observed when the subject was inanimate. The second hierarchy is the weight hierarchy, operationalised in terms of the length of the subject NP, measured in number of

words.[3] The rate of OS order increased with increasing subject length, from about 55% for subjects consisting of a single word to 79% for subjects made up of 10 words. The animacy hierarchy and the weight hierarchy were selected for the following reasons: first, they represent different types of information; second, they had a strong influence on the frequencies of SO and OS order; and third, they do not depend on a prior context, which makes them particularly appropriate for experiments which present isolated sentences.

In Experiment 1, the subject consists of either two or four words. An example item from this experiment is shown with an animate subject in (4) and with an inanimate subject in (5). For reasons of space, only OS order is shown. Subject length was varied by including or omitting the two-word adjective phrase in parentheses. Thus, the subject, which was always a definite NP, consisted of either two or four words. The verb was always an object-experiencer psych verb because such verbs easily take animate and inanimate subjects.

(4) Der Reporter hat gesagt, dass ihn der (äußerst erfolgreiche) Stürmer
the reporter has said that him the extremely successful striker
fasziniert hat.
fascinated has
'The reporter said that the (extremely exciting) striker fascinated him.'

(5) Der Reporter hat gesagt, dass ihn das (äußerst spannende) Endspiel
the reporter has said that him the extremely exciting) final
fasziniert hat.
fascinated has
'The reporter said that the (extremely exciting) final fascinated him.'

For sentences as investigated in the upcoming Experiment 1, the left part of Table 1 shows the percentages of OS order in the corpus study of Bader (2020). The corpus data include sentences with verbs of all kinds because otherwise, the number of observations would have been too low. For this reason, the rate of OS order may be underestimated because object-experiencer verbs belong to the class of verbs that show a preference for OS order. To address this issue, a production experiment was run using the exactly same materials as Experiment 1 (Bader, in preparation); thus, all sentences contained an object experiencer verb. The percentages of sentences produced with OS order in this production experiment are shown in the right part of Table 1. Participants first read a main clause like *Der Opa hat den Enkel beeindruckt* 'The grandpa impressed the grandson'. They then had to reproduce the main clause from memory, following a prompt like *Der*

3. The weight hierarchy is taken here as a shorthand for more refined accounts of how weight affects word order; cf. Hawkins (2004), among others.

Enkel hat gesagt, dass 'The grandson said that'. In order to fit the prompt, the main clause had to be transformed into an embedded clause containing an object pronoun, which could either precede or follow the non-pronominal subject. Thus, a main clause like *Der Opa hat den Enkel beeindruckt* could either be transformed to *dass der Opa ihn beeindruckt hat* or to *dass ihn der Opa beeindruckt hat*, both meaning 'that the grandpa impressed him'. The initial main clause and the prompt had to be read aloud. The embedded clause transformed from the main clause was produced orally.

An inspection of Table 1 shows effects of subject animacy and subject length in both the corpus data and the experimental data, with two differences. First, the rate of OS order is about 10% lower in the corpus data than in the experimental data, which can be attributed to the inclusion of verbs of all sorts in the corpus study but only object-experiencer verbs in the experimental study. Second, an additive pattern is visible in the corpus data, with an increase of about 16.5% when going from sentences with two-word subjects to sentences with four-word subjects; but an interactive pattern is visible in the experimental data, where an increase of similar size is seen when the subject is animate, whereas there is basically no length effect for sentences with inanimate subjects. Given the already high rate of OS order with short inanimate subjects, the absence of a further length effect may be a ceiling effect. How these frequency differences are reflected in acceptability ratings is tested in the next section.

Table 1. Percentages of OS order depending on the animacy and length of the subject NP. Corpus data are from Bader (2020) and experimental data from Bader (in preparation). The factors animacy and length of subject NP correspond to the conditions of Experiment 1. Note: Corpus data include both dative and accusative pronouns, experimental data include accusative pronouns only

	Corpus data		**Experiment**	
	Animate subj	**Inanimate subj**	**Animate subj**	**Inanimate subj**
Short subj (2 words)	40	60	50	82
Long subj (4 words)	56	77	67	83

3. Experiments 1–3: How the position of an object pronoun affects sentence acceptability

This section presents three experiments that have investigated the acceptability of sentences with an object pronoun either before or after a nonpronominal sub-

ject using the method of magnitude estimation. This method, which goes back to work in psychophysics by Stevens (1957), has been adapted for linguistic purposes by Bard, Robertson, and Sorace (1996) and Cowart (1997). Magnitude estimation allows participants to rate sentences on an open-ended, continuous numerical scale and can, therefore, uncover fine distinctions in acceptability. When magnitude estimation was introduced to linguistics, this was seen as a distinctive advantage of the method. Later research showed that ratings procedures that make use of discrete rating categories – Likert scales which typically range from 1–5 or 1–7, and even the binary distinction between grammatical and ungrammatical sentences – deliver acceptability measures of comparable quality (e.g., Bader and Häussler 2010a; Weskott and Fanselow 2011; see overviews in Featherston 2021; Goodall 2021). The reason for running the experiments reported in this paper using the magnitude estimation procedure was therefore a purely practical one. Because many experiments in our lab used magnitude estimation at the time when the research reported here was conducted, it was most convenient to also use this method.[4]

Experiment 1 probes how animacy and length of the subject NP affect the acceptability of object pronoun sentences with either SO or OS word order. It reveals some variation in acceptability, but overall, all sentences receive rather high acceptability ratings. In order to ascertain that the overall high acceptability observed in Experiment 1 is not due to participants being insensitive to constraints on pronoun position, Experiment 2 tests the generalisation that pronominal objects must not occur later than directly after a clause-initial subject. The final Experiment 3 investigates the effect of the subject's length more closely.

3.1 Experiment 1

As illustrated above in (4) and (5), Experiment 1 investigates two major properties of non-pronominal subjects that have been found to affect the placement of object pronouns before or after the subject – animacy and length. In all sentences investigated in Experiment 1, the object pronoun is contained in an embedded clause. The subject of this clause is a definite NP which is either animate (human) or inanimate and two or four words long (for an original experimental sentence, see (4) and (5)). In order to hold the verb constant across conditions, the verb of the embedded clause is always an object experiencer verb like *erfreuen* 'please' because these verbs allow both animate and inanimate subjects.

4. Experiment 1 has recently been replicated using ratings on a 1–7 scale, with by and large the same results.

As pointed out in the introduction, the relationship between corpus frequencies and acceptability ratings is a controversial issue, which makes it difficult to derive definite predictions. For the sake of the argument, let us make the strongest assumption according to which acceptability ratings mirror corpus frequencies in a direct manner (see Bresnan et al. 2007, for a proposal along this line). Under this assumption, the production data shown in Table 1 predict main effects of animacy and length modified by an interaction between these two factors. With a short animate subject, SO and OS order are not far apart in terms of frequency, so both should be rated as equally acceptable. When the subject gets longer or is inanimate, OS order outweighs SO order, so OS order should become more acceptable than SO order. According to the frequency data, the factors have additive effects; according to the experimental production data, they interact, so that acceptability may not go up further when an OS sentence contains a subject that is long and inanimate. However, when the interaction seen for the experimental data is due to a ceiling effect, as indicated above, the two factors should affect acceptability in an additive way because the rating scale used by magnitude estimation is an open ended scale (at least in principle; see Ellsiepen and Bader 2014 for discussion).

3.1.1 *Method*

Participants

52 students at Goethe University Frankfurt participated in Experiment 1. In this and the following two experiments, all participants were native speakers of German and naive with respect to the purpose of the experiment. Participants attended an introductory psycholinguistics course unrelated to the current experiments and received course credit for participation.

Table 2. Sample materials for Experiment 1

Inanimate	
SO	Der Opa hat gesagt, dass das (äußerst wertvolle) Buch ihn erfreut hat.
	the grandpa has said that the extremely valuable book him pleased has
OS	Der Opa hat gesagt, dass ihn das (äußerst wertvolle) Buch erfreut hat.
	the grandpa has said that him the extremely valuable book pleased has
'Grandpa said that the extremly valuable book had pleased him.'	
Animate	
SO	Der Opa hat gesagt, dass der (äußerst lustige) Enkel ihn erfreut hat.
	the grandpa has said that the extremely funny grandson him pleased has
OS	Der Opa hat gesagt, dass ihn der (äußerst lustige) Enkel erfreut hat.
	the grandpa has said that him the extremely funny grandson pleased has
'Grandpa said that the extremly funny grandson had pleased him.'	

Materials

48 sentences were constructed for Experiment 1. A complete stimulus set is shown in Table 2. Each sentence consisted of a main clause followed by an embedded clause. The main clause always started with a definite subject NP denoting a person that served as antecedent for the upcoming object pronoun. The main clause subject was followed by the verb phrase *hat gesagt* "has said" and an embedded *that*-clause. The *that*-clause in turn consisted of a definite subject NP, an object pronoun, and one of the object experiencer verbs in (6) in the perfect tense. Each verb was used in four sentences.

(6) *erfreuen* 'to please', *interessieren* 'to interest', *erstaunen* 'to astonisch', *beeindrucken* 'to impress', *stören* 'to disturb', *ärgern* 'to bother', *überraschen* 'to surprise', *enttäuschen* 'to disappoint', *erheitern* 'to amuse', *faszinieren* 'to fascinate', *schockieren* 'to shock', *verwirren* 'to baffle'

For each sentence, eight versions were created according to the three two-way factors *Animacy, Length* and *Order*. All factors applied to the embedded clause. The factor Animacy varied the animacy of the subject NP, which was either animate (human) or inanimate. The factor Length varied the length of the subject NP. A short subject was a two-word NP consisting of a definite determiner and a noun; a long subject was a four-word NP containing an adverb and an adjective between the determiner and the noun. The final factor Order manipulated the order of subject and object pronoun, which was either subject-before-object (SO) or object-before-subject (OS).

The 48 sentences were distributed onto eight lists according to a Latin Square design. Each list contained one version of each sentence and an equal number of sentences in each condition. The experimental lists were individually randomised for each participant and then interspersed in a list of filler sentences of varying structures. Different filler lists were used, containing between 50 and 75 fillers. The majority of the filler sentences was from unrelated experiments. Each participant saw only one list.

Procedure

The ME procedure used in the following three experiments closely followed the procedure described in Bard, Robertson, and Sorace (1996); Sorace (2000) and Keller (2000). Each experimental session consisted of three phases, which were run using software developed by myself. In the customisation phase, participants were acquainted with the principles of ME by judging the length of six lines presented on a computer screen. In the training phase, they judged the acceptability of ten training sentences. In the final experimental phase, they judged the acceptability of the experimental and filler sentences described in the materials section.

Each phase consisted of the following steps. First, participants read an instruction that was displayed on the computer screen. The instruction explained the ME procedure with the help of an example. At the end, the main points were repeated in the form of a list. Participants were encouraged to contact the experimenter in case there were any questions regarding the task. When participants had finished reading the instruction, they pressed a button which triggered the display of the reference stimulus (either a line or a sentence). Participants assigned a numerical value to the reference stimulus. Afterwards, the experimental stimuli were displayed one by one, and participants judged each stimulus relative to the reference stimulus. The reference stimulus, as well as the reference value assigned to it, remained visible on the computer screen while participants worked through the experimental stimuli. Participants typed their judgements using a regular keyboard. Judgements and judgement times were recorded automatically.

To enhance comparability with prior work, the reference sentence for the final experimental phase was taken almost literally from Keller (2000, sentence (B.18)/page 377). As shown in (7), the reference sentence is a sentence with a definite inanimate object preceding a definite animate subject. Such sentences (so called 'scrambling' sentences) are grammatical but of reduced acceptability, at least when presented out of context (cf. Keller 2000; Bader and Häussler 2010a).

(7) Ich glaube, dass den Bericht der Chef in seinem Büro gelesen hat.
I believe that the-ACC report the-NOM boss in his office read has
'I believe that the boss read the report in his office.'

Scoring

All statistical analyses reported in this paper were conducted using the R statistics software (R Core Team 2020). The acceptability data were analysed with linear mixed-effects models (Bates et al. 2015b). Models included the experimental factors and their interactions as fixed effects. Random effect terms were determined following the model fitting procedure proposed in Bates et al. (2015a). Fixed effects were entered into the model using effect coding (0.5 vs −0.5 in the case of two-level factors), that is, the intercept represents the unweighted grand mean and fixed effects compare factor levels to each other. Where necessary, simple contrasts were computed to compare mean values.

Computing the results for an ME experiment is somewhat involved because the numerical scores that participants assign to the experimental sentences can only be interpreted in relation to the reference value that each participant gives to the reference sentence at the beginning of the experiment. For example, if an experimental sentence is assigned a value of 100, the meaning of this value is very different depending on whether the initial reference sentence received a value of

50 (in which case the experimental sentence was judged as being twice as good as the reference sentence) or a value of 200 (in which case the experimental sentence was judged as being only half as good as the reference sentence). Therefore, the raw judgements obtained in an ME experiment have to be transformed in such a way that the value assigned to the reference sentence is taken into account. In the following, a commonly used transformation is employed which scales the individual scores of each participant directly in relation to the participant's reference value. This is achieved by dividing the score assigned to each experimental sentence by the value assigned to the reference sentence. In order to approximate a normal distribution, the resulting ratio is further transformed by taking the logarithm to base ten (cf. Bard, Robertson, and Sorace 1996).

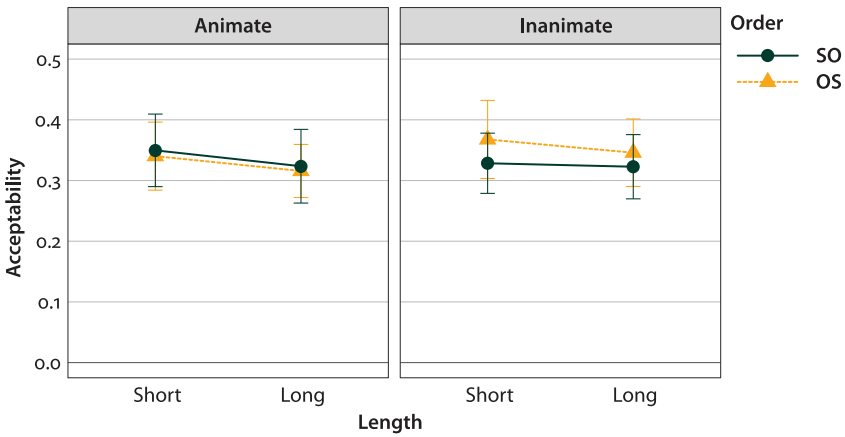


Figure 1. Mean acceptability ratings in Experiment 1

Table 3. Mixed-effects model for acceptability ratings in Experiment 1.
Formula: *response ~ Animacy*Length*Order + (1+length||subject) + (1+length||sentence)*

Contrast	Estimate	Std. Error	Df	T value	> |z|
Intercept	0.337	0.055	55	6.08	$p < 0.01$
Animacy	−0.010	0.007	2294	−1.33	n.s.
Length	0.020	0.010	45	1.93	$p = 0.060$
Order	−0.012	0.007	2298	−1.86	$p = 0.064$
Animacy × Length	0.012	0.013	2294	0.86	n.s
Animacy × Order	0.041	0.013	2298	3.03	$p < 0.01$
Length × Order	−0.007	0.013	2302	−0.56	n.s.
Animacy × Length × Order	0.018	0.027	2302	0.68	n.s.

3.1.2 *Results*

Figure 1 shows the mean acceptability ratings obtained in Experiment 1. The zero point on the y-axis in Figure 1 means 'of equal acceptability as the reference sentence'. The corresponding linear mixed-effects model is summarised in Table 3. Overall, the sentences of Experiment 1 received a mean acceptability rating of about 0.34, which is significantly higher than zero, as shown by the significant intercept in Table 3. Thus, on average, the sentences of Experiment 1 were judged as twice as acceptable as the reference sentence in (7), an OS sentence with a full definite NP preceding the subject. In addition, Table 3 reveals two marginally significant main effects and a significant interaction.[5] Sentences with a short subject were judged as slightly more acceptable than sentences with a long subject, resulting in a marginally significant main effect for Length. Since Length did not interact with any other factor, this means that sentences with a long subject appeared somewhat less acceptable to participants than sentences with a short subject, independently of the order of subject and object. The main effect of Order also reached marginal significance, but this effect has to be qualified by a significant interaction with Animacy. This interaction reflects the finding that Animacy had no effect on SO sentences (0.34 versus 0.33; $t = 1.20$, $p > 0.1$), whereas for OS sentences, an inanimate subject led to a somewhat higher acceptability than an animate subject (0.36 versus 0.33; $t = 3.08$, $p < 0.01$).

The mean judgements time was 6,051 ms, where judgement times include reading the sentence, determining an acceptability value, and typing the value into a field on the computer screen. Judgement times were also analysed statistically, but for reasons of space, the analysis is reported only in a cursory way. In addition to a main effect of length (5,532 ms for sentences with short subject versus 6570 ms for sentences with long subjects), a significant interaction between Order and Animacy was found which mirrors the interaction for the acceptability ratings. Whereas judgement times for SO sentences were not affected by Animacy (6,028 ms with animate subject versus 6,032 ms with inaminate subject), a significant difference became evident for OS sentences (6,282 ms with animate subject versus 5,862 ms with inaminate subject)

5. Following standard practice, p-values between 0.05 and 0.10 are considered marginally significant in this paper; for a critical discussion of the pro and cons of reporting marginally significant results, see Olsson-Collentine, Van Assen, and Hartgerink (2019) and references cited there.

3.1.3 *Discussion*

Experiment 1 has three main findings. First, acceptability was high in all conditions, with a mean value of 0.337 on the logarithmic scale. This means that, on average, the sentences of Experiment 1 were judged as being twice as acceptable as the reference sentence in (7) (a sentence with a scrambled definite object). Bader and Häussler (2010a) used the same reference sentence in their magnitude estimation experiments and obtained binary judgement data in addition to magnitude estimation ratings. Sentence with ratings above 0.3 on the logarithmic scale were judged as grammatical in the binary judgements task with about 90% of the time. This allows the conclusion that all sentences investigated in Experiment 1 were perceived as fully grammatical by the participants. Second, Animacy interacted with Word Order in the way expected given what has been found in corpus studies: SO and OS sentences were of equal acceptability when the subject was animate, but with an inanimate subject, OS sentences were somewhat more acceptable than SO sentences. As pointed out in the introduction, whether an observed acceptability difference should be attributed to the grammar or the processing mechanisms is often not easy to decide, and this is especially true for relatively small differences between sentences that are of overall high acceptability. A further discussion of this finding will, therefore, be postponed to the general discussion when all data have been presented. Third, in contrast to Animacy, Length did not interact with Order but only showed a marginally significant main effect, indicating that sentences with long subjects were somewhat less acceptable than sentences with short subjects, independently of the order between subject and object pronoun. This contrasts with corpus findings showing that the rate of OS order increases with increasing length of the subject NP for the length manipulation applied in Experiment 1 (2 versus 4 words). This adds to the existing evidence that highly acceptable syntactic variants can differ in terms of frequency without a related difference in terms of acceptability (see Arppe and Järvikivi 2007; Bader and Häussler 2010a), but it does not preclude that a stronger length manipulation leads to acceptability differences between SO and OS sentences. This possibility is explored in Experiment 3.

3.2 Experiment 2

In order to better appreciate the relatively small acceptability differences found in Experiment 1, Experiment 2 investigates the grammatical constraint that nothing else than a subject NP can separate a weak object pronoun from the left edge of the middlefield, as captured in the template for the middlefield of German sentences in (3). This template makes the prediction that the word order freedom

enjoyed by object pronouns is confined to a small part of the middlefield – an object pronoun must appear directly after the complementiser or, if the subject immediately follows the complementiser, directly after the subject. Later positions, in contrast, should lead to a strong decrease in acceptability.

In order to test this prediction, Experiment 2 varies the position of an adverbial that occurs in addition to subject and object. As shown in (8), an adverbial can appear in one of three positions relative to subject and object pronoun – preceding both subject and object ($ADV_{initial}$), between subject and object (ADV_{medial}), and following subject and object (ADV_{final}).

(8) a. dass *$ADV_{initial}$ *subject* *ADV_{medial} **pro** ADV_{final}… verb.
a′ dass (heute) *der Bürgermeister* (heute) **ihn** (heute) besucht.
that today the mayor today him today visits
'that the mayor is visiting him today.'
b. dass *$ADV_{initial}$ **pro** ADV_{medial} *subject* ADV_{final}… verb.
b′ dass (heute) **ihn** (heute) *der Bürgermeister* (heute) besucht.
that today him today the mayor today visits
'that the mayor is visiting him today.'

The starred adverbial positions in (8) are those that are banned according to the template in (3). If this template correctly captures the grammatical knowledge of speakers of German, the predictions listed below follow. These predictions are put to an empirical test in Experiment 2.

- Adverbial in final position:
  Acceptability should be high for both SO-Adv and OS-Adv order because the pronoun precedes the adverbial in both cases.
- Adverbial in medial position:
  Acceptability should be high for O-Adv-S order where the pronoun precedes the adverbial but low for S-Adv-O order where the pronoun follows the adverbial.
- Adverbial in initial position:
  Acceptability should be low for both Adv-SO and Adv-OS order because the pronoun follows the adverbial in both cases.

The constraint that weak object pronouns have to appear in front of adverbials in the German middlefield is probably a consequence of a more general ordering principle according to which given and backgrounded elements should appear in early positions in the middlefield whereas new and focused elements should appear in late positions (e.a. Lenerz 1977; Diesing 1992). Like in Experiment 1, the subject NP is always a definite NP in the sentences investigated in Experiment 2 (see Table 4). In contrast to weak object pronouns, definite NPs are allowed to

either precede or follow the adverbial, but their preferred position is in front of the adverbial. This predicts that acceptability should be lower when subjects follow the adverbial, although not to the same extent as expected for object pronouns.

Table 4. Sample materials for Experiment 2

Final adverbial position	
SO	Der Opa hat gesagt, dass das Buch ihn **schon** erfreut hat.
	the grandpa has said that the book him indeed pleased has
OS	Der Opa hat gesagt, dass ihn das Buch **schon** erfreut hat.
	the grandpa has said that him the book indeed pleased has
Middle adverbial position	
SO	Der Opa hat gesagt, dass das Buch **schon** ihn erfreut hat.
	the grandpa has said that the book indeed him pleased has
OS	Der Opa hat gesagt, dass ihn **schon** das Buch erfreut hat.
	the grandpa has said that him indeed the book pleased has
Initial adverbial position	
SO	Der Opa hat gesagt, dass **schon** das Buch ihn erfreut hat.
	the grandpa has said that indeed the book him pleased has
OS	Der Opa hat gesagt, dass **schon** ihn das Buch erfreut hat.
	the grandpa has said that indeed him the book pleased has

3.2.1 *Method*

Participants

42 students at Goethe University Frankfurt participated in Experiment 2. None of the participants had already participated in Experiment 1.

Materials

30 sentences from Experiment 1 were modified in the following way to serve as materials for Experiment 2 (see Table 4 for a complete stimulus example). First, only the versions with a short inanimate subject were retained. Second, a modal adverb/particle (see (9) for the full list) was inserted into the embedded clause at one of three possible positions: before both subject and object, between subject and object, and after subject and object. Thus, each sentence in Experiment 2 occurred in six conditions according to the factors Order (SO or OS) and Adverb Position (initial or medial or final).

(9) *schon* 'already', *doch* 'after all', *eigentlich* 'rather', *aber* 'however', *eben* 'just'

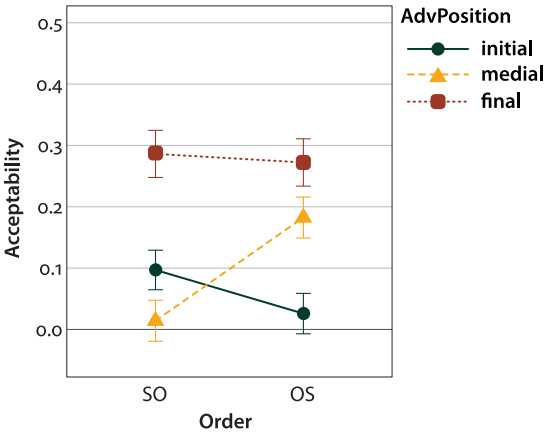


Figure 2. Mean acceptability ratings in Experiment 2

Table 5. Mixed-effects model for acceptability ratings in Experiment 2.
Formula: *response ~ AdvPosition*Order + (1|participant) + (1|sentence)*

Contrast	Estimate	Std. Error	Df	T value	> |z|
Intercept	0.146	0.032	43	4.64	< 0.01
AdvPosition1	0.181	0.012	1184	14.98	< 0.01
AdvPosition2	0.037	0.012	1184	3.06	< 0.01
Order	−0.028	0.010	1184	−2.82	< 0.01
AdvPositionl × Order	0.182	0.024	1184	7.55	< 0.01
AdvPosition2 × Order	−0.239	0.024	1184	−9.91	< 0.01

Note. AdvPositionl = final vs. medial; AdvPosition2 = medial vs. initial

Procedure

Experiment 2 used the same procedure as Experiment 1.

3.2.2 *Results*

Figure 2 shows the mean ratings for the sentences tested in Experiment 2. For the three-level factor Adverb Position, two pairwise contrasts were defined, one comparing the level with the highest rating to the level with the intermediate rating (final versus medial) and one comparing the level with the intermediate rating to the level with the lowest rating (medial versus initial). Table 5 shows the corresponding linear mixed-effects model. All main effects as well as all interactions turned out to be significant. As can be seen in Figure 2, the relationship between SO and OS order is different in each of the three adverbial conditions.

(i) With a final adverbial and thus both subject and object preceding the adverbial, acceptability is high and there is no significant difference between SO and OS (0.29 versus 0.27, $t = 0.82$, n.s.). (ii) When the adverbial occurs between subject and object, it precedes the pronoun with SO order but follows it with OS order. With this adverbial position, acceptability is high for OS order but low for SO order (0.18 versus 0.01, $t = 9.86$, $p < 0.01$). As also shown in Figure 2, although the pronoun in OS sentences precedes both a final and a medial adverbial, acceptability is higher with a final than with a medial adverbial (0.27 versus 0.18, $t = 5.26$, $p < 0.01$). Thus, as predicted, acceptability decreases when the subject follows the adverbial, but the reduction is substantially smaller than in the case of a pronoun following the adverbial. (iii) An initial adverbial precedes both subject and object, which should lead to reduced acceptability. As can be seen in Figure 2, this prediction is borne out, but the reduction in acceptability is less pronounced for SO than for OS sentences (.09 versus 0.03, $t = 4.16$, $p < 0.01$).

The judgement time analysis revealed a significant effect of word order on sentences with a medial adverbial with longer judgement times for OS than for SO sentences. At first glance, this seems counter-intuitive, because judgements times are higher for sentences with higher acceptability ratings than for sentences with lower acceptability ratings. However, as shown by Bader and Haussier (2010a), judgement times and acceptability ratings are not monotonically related to each other in the magnitude estimation task. Instead, judgement times are typically fastest for low and high acceptability values and slowest for acceptability values in the middle range. This is in accordance with the finding of faster judgement times for SO sentences with medial adverbial (low acceptability) than for OS sentences with medial adverbial (mid-level acceptability).

3.2.3 *Discussion*

The major result yielded by Experiment 2 is the strong drop in acceptability caused by putting an object pronoun behind an adverbial. When the object pronoun preceded the adverbial, sentences were judged as being about twice as acceptable as the reference sentence (unless the subject followed the adverbial, leading to an independent decrease in acceptability). This replicates the findings from Experiment 1 for sentences without an adverbial. When the pronoun immediately followed the adverbial, in contrast, sentences were judged as of about equal acceptability as the reference sentence. That is, a sentence with a pronoun after an adverbial is of about equal acceptability to a scrambling sentence out of context. Such sentences are, thus, not plainly ungrammatical but heavily marked. In particular, given that sentences were presented visually, readers may have assigned an implicit prosody with an accent on the pronoun, thereby making the pronoun strong so that it was exempted from the positional constraints on weak object pronouns.

An additional finding of Experiment 2 was that sentences in which the subject followed the adverbial were also reduced in acceptability, but the reduction was not as strong as the one found for the pronominal object. In comparison to sentences with a final adverbial, which received a mean acceptability rating of about 0.28 and can thus be taken as a fully acceptable baseline, putting the subject behind the adverbial (order O-Adv-S) resulted in an acceptability value of 0.18, but putting the object behind the adverbial (order S-Adv-O) in a much lower acceptability value of 0.01.

In summary, the results yielded by Experiment 2 confirm the prediction that an object pronoun has two acceptable positions: directly after the complementiser or, in cases where the subject immediately follows the complementiser, directly after the subject. Later positions, in contrast, lead to a strong decrease in acceptability. The finding that all sentences in Experiment 1 received high acceptability ratings, with only small variation due to animacy, does not, therefore, reflect a general insensitivity of native speakers with regard to the position of object pronouns in the middlefield. Instead, the high acceptability values in all conditions of Experiment 1 must be attributed to the fact that in the first experiment, the object pronoun always occurred in one of the two positions reserved for weak object pronouns.

3.3 Experiment 3

In Experiment 1, the weight of the subject NP, operationalised in terms of length measured in number of words, did not affect SO and OS sentences in different ways, as one could have expected given the effect that weight has on production frequencies. Experiment 3 tests whether a stronger length manipulation leads to acceptability differences between SO and OS sentences. To this end, Experiment 3 investigates subject NPs of three different lengths, as shown in (10) (for reasons of space, only the SO variant of each sentence is shown).

(10) a. Der Opa hat gesagt, dass *das Buch* ihn erfreut hat.
the grandpa has said that the book him pleased has
'Grandpa said that the book pleased him.'

b. Der Opa hat gesagt, dass *das äußerst wertvolle Buch* ihn erfreut hat.
the grandpa has said that the extremely valuable book him pleased has
'Grandpa said that the extremely valuable book pleased him.'

c. Der Opa hat gesagt, dass *das Buch*, *das äußerst wertvoll war*, ihn erfreut hat.
the grandpa has said that the book which extremely valuable was him pleased has
'Grandpa said that the book which had been extremely valuable pleased him.'

The subjects in (10a) and (10b) are identical to the subjects in the short and long condition of Experiment 1 and thus contain two and four words, respectively. The subject in the new condition in (10c) is a definite NP modified by a relative clause consisting of four words, for a total subject length of six words. Lengthening the subject NP by means of a relative clause does increase its weight not only in terms of number of words, but also in terms of syntactic and prosodic structure. Although different measures of phrasal weight are heavily confounded (see Wasow 2002), increasing weight in several dimensions guarantees that there is a substantial weight increase when going from two-word subjects to six-word subjects.

3.3.1 *Method*

Participants
101 students at Goethe University Frankfurt participated in Experiment 3. None of the participants had participated in either Experiment 1 or 2.

Materials
For Experiment 3, 30 sentences of Experiment 1 were modified as follows. First, all sentence versions with an animate subject in the embedded clause were removed, leaving only inanimate subjects. Second, a third level was added to the originally two-way factor Length. Besides subject consisting of either two (Det N) or four words (Det Adv A N), subjects consisting of six words were created by embedding the adverb and the adjective of the four-word condition into a relative clause following the noun. In addition to the adverb and the adjective, the relative clause contained an initial relative pronoun and a final finite copula. The relative clause was thus always made up of four words, which together with the definite article and the head noun resulted in a total subject length of six words. As a result of the two changes made to the materials of Experiment 1, each sentence of Experiment 3 appeared in six versions according to the two factors Order (SO or OS) and Length (two, four or six words).

Procedure
The same procedure was used as in the preceding experiments.

3.3.2 *Results*

Figure 3 shows the mean acceptability ratings obtained in Experiment 3. Factor coding followed the procedures of the preceding experiments. The corresponding mixed-effects model is summarised in Table 6. As in all conditions of Experiment 1 and in Experiment 2 with an adverbial in final position, acceptability is relatively high throughout. Also as in Experiment 1, Figure 3 shows a small decline in acceptability with increasing length. In contrast to Experiment 1, however, acceptability

decreases with length in a more pronounced way for SO than for OS sentences. As a result of this, SO sentences are judged somewhat less acceptable than OS sentences for longer subjects (four and six words), whereas acceptability for SO and OS sentence is almost identical for short subjects (two words).

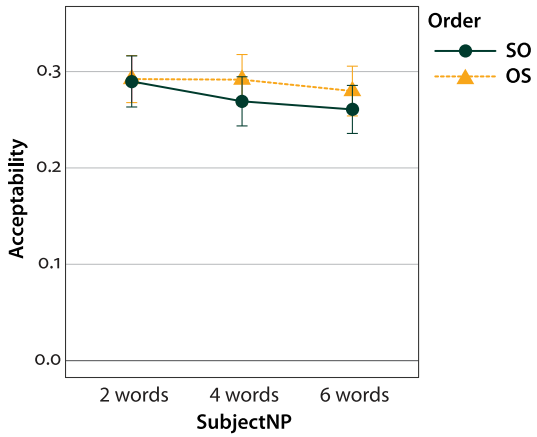


Figure 3. Mean acceptability ratings in Experiment 3

Table 6. Mixed-effects model for acceptability ratings in Experiment 3.
Formula: *response ~ SubjectNP*Order + (1|participant) + (1|sentence)*

Contrast	Estimate	Std. Error	Df	T value	> |z|
Intercept	0.281	0.025	104	11.24	< 0.01
SubjectNP1	0.011	0.006	2895	1.74	0.082
SubjectNP2	0.010	0.006	2895	1.66	0.098
Order	−0.015	0.005	2895	−2.95	< 0.01
SubjectNP1 × Order	0.020	0.012	2895	1.63	0.10
SubjectNP2 × Order	−0.003	0.012	2895	−0.27	0.79

Note. SubjectNP1 = 2 words vs. 4 words; SubjectNP2 = 4 words vs. 6 words

As shown in Table 6, only the main effect of Order was fully significant. The small drops in acceptability caused by lengthening the subject NP, in contrast, only resulted in a marginally significant effect. Furthermore, the interaction visible in Figure 3 between Order and the first two levels of Length (two words versus four words) even failed to reach marginal significance. Pairwise comparisons testing the effect of Order separately for each level of Length, however, revealed a difference between sentences with a two-word subject and sentences with either four- or six-word subjects. Order had no effect on sentences with a 2-word subject

(0.30 versus 0.29; $t = 0.29$, $p > 0.1$) but on sentences with a four-word subject (0.27 versus 0.29; $t = 2.60$, $p > 0.1$) and on sentences with a six-word subject (0.26 versus 0.28; $t = 2.22$, $p > 0.1$).

The analysis of the judgement times for Experiment 3 showed only significant effects of length, with faster judgements times for sentences with 2-word subject than for sentences with 4-word subjects, and faster judgements times for sentences with 4-word subject than for sentences with 6-word subjects.

3.3.3 *Discussion*

Experiment 3 provides weak evidence that the weight-based ordering principle 'short before long' (that is, Behagel's fourth law) affects the perception of sentences as more or less acceptable. With long subject NPs (four- or six-word subjects), acceptability was lower for sentences with SO order, which are dispreferred according to this law, than sentences with OS order. The evidence is only weak, however, for two reasons. First, the observed effects were small. Second, the statistical analysis reached full significance only in the pairwise comparisons but not in the full model. This is in contrast to findings from language production, for which robust effects of length have been found both in corpus studies (Heylen 2005; Bader 2020) and in experimental investigations (Bader, in preparation). Possible conclusions to be drawn from the observed discrepancy between accessibility and production data are considered in the general discussion.

4. General discussion

The starting point of this paper was the observation that the order of a non-pronominal subject and a pronominal object within the German middlefield seems to be a case of free variation. That is, even when sentences are matched with regard to features that are known to affect word order – in particular, features related to lexical semantics (e.g. animacy, thematic roles), discourse structure (e.g. givenness) and weight (e.g. number of words) – still both orders are produced, as revealed by corpus studies. However, as also revealed by corpus analyses, the odds of selecting either one or the other of the two possible orders varies strongly depending on the particular combination of features. This raises the question of whether the observed differences in production frequencies are mirrored by corresponding differences in acceptability. This paper has presented three experiments that addressed this question.

The experiments yielded the following main findings. First, acceptability was always high when the object pronoun appeared in the first position of the middlefield or the subject occurred in the first position and the object pronoun

immediately thereafter. In contrast, when the object pronoun occupied a later position, as it did in some conditions of Experiment 2, acceptability was substantially reduced. This finding is in accordance with the corpus-based word order template of the German middlefield given in (3). Second, although acceptability was generally high when the object pronoun occurred in one of the two positions allotted to it in the template in (3), some variation of acceptability depending on the animacy and length of the subject was still found. In Experiment 1, a small order-independent decrease in acceptability for sentences with longer subjects was observed, as was a small increase in acceptability for OS sentences with an inanimate subject in comparison to OS sentences with an animate subject. In Experiment 3, acceptability again decreased somewhat with increasing subject length, but in addition there were indications that the decrease was more pronounced for SO than for OS sentences.

As pointed out in the introduction, a difference in acceptability can be caused by differences in grammaticality, by processing differences, or by a combination of both. In the remainder of this paper, I will argue that the small acceptability differences revealed by Experiments 1 and 3 should not be attributed to the grammar but to the processing systems. There are two main reasons for this hypothesis. First, the observed differences were small, especially when compared to the drop in acceptability when an object pronoun followed an adverbial in Experiment 2. For example, in Experiment 1, a significant difference between sentences with animate and inanimate subjects was observed for OS sentences: inanimate subjects resulted in a rating of 0.357 on the logarithmic scale, but animate subjects resulted in a rating of only 0.328. This contrasts with a drop in Experiment 2 from 0.272 for OS sentences with a final adverbial to 0.026 for OS sentences with an initial adverbial. Furthermore, the effect size due to word order for sentences with an inanimate subject is about the same as the order-independent decrease in acceptability for sentences with long subjects in Experiment 1, which was from 0.346 for sentences with a short subject to 0.327 for sentences with a long subject. Second, the observed differences were not fully reliable. For example, the significant difference between SO and OS sentences with an inanimate subject was not replicated in Experiments 2 and 3. Furthermore, while Experiment 1 revealed only an order-independent effect of length, Experiment 3 found that acceptability decreased with increasing subject length more for SO than for OS sentences.

Given the widespread recognition that grammaticality itself is not a binary property but comes in grades that reflect weighted constraints (see Goodall 2021, and references cited there), small differences in acceptability do not per se exclude an account in terms of grammar. However, together with the fleeting nature of the differences observed in Experiments 1–3, it seems unlikely that we are dealing with differences encoded in the grammar. Consider, for example, the acceptabil-

ity differences related to length. Effects of this kind are a classic case for an explanation in terms of processing mechanisms (e.g. Gibson 2000). Given that both order-dependent and order-independent effects of length were found, it is most parsimonious to attribute all length effects to the processing mechanisms.

In conclusion, the experimental results reported in this paper support the claim that the relative order of non-pronominal subject and pronominal object is truly a matter of free variation within the grammar. The grammar requires object pronouns to occur at the left edge of the middlefield, or, when the subject occupies the first position of the middlefield, directly thereafter, but the grammar favours neither of these two positions. Thus, whether an object pronoun occurs in middlefield-initial position or immediately after a midfield-initial subject does not have any bearing on meaning or acceptability. It is, therefore, left to the language production mechanisms to decide in which of the two positions allowed by the grammar an object pronoun is produced. Since this decision is subject to the usual probabilistic constraints on linearisation, the order of object pronoun and non-pronominal subject observed in language production is a matter of free, but not random, variation.

A question that cannot be answered from the currently available evidence is whether our language production mechanisms select a position for an object pronoun in the middlefield in a non-random or in a deterministic way. For example, in the corpus study of Bader (2020), a logistic regression model based on a number of word-order hierarchies predicted the observed order in 76.7% of all cases. By taking factors into account that were not included in this corpus study – for example, factors related to the preceding context and factors related to individual properties of writers – this value can likely be improved, but will it approach 100% or will a certain random element remain, even if all relevant factors have been taken into account? My guess is that a certain amount of randomness will remain. This would be in line with other cognitive processes for which it is commonly assumed that choices are modulated to some degree by random noise (e.g. Anderson 2009).

References

Anderson, John R. 2009. *How can the human mind occur in the physical universe?* Oxford: Oxford University Press.

Arppe, Antti & Järvikivi, Juhani. 2007. Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3(2): 131–159.

Bader, Markus. 2020. The position of object pronouns in the German midfield. *Linguistics* 58(4): 1059–1115.

Bader, Markus. in preparation. *Conceptual accessibility and weight in sentence production: The case of German object pronouns*. Manuscript in preparation.

Bader, Markus & Häussler, Jana. 2010a. Toward a model of grammaticality judgments. *Journal of Linguistics* 46(2): 273–330.

Bader, Markus & Häussler, Jana. 2010b. Word order in German: A corpus study. *Lingua* 120(3): 717–762.

Bard, Ellen G., Robertson, Dan & Sorace, Antonella. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1): 32–68.

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation Journal* 23(3). 209–226.

Bates, Douglas, Kliegl, Reinhold, Vasishth, Shravan & Baayen, R. Harald. 2015a. Parsimonious mixed models. *arXiv.org preprint – arXiv:1506.04967 [stat.ME]*. https://arxiv.org/ abs/1506 .04967v1 (20 February 2023).

Bates, Douglas, Mächler, Martin, Bolker, Ben & Walker, Steve. 2015b. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48.

Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana & Baayen, R. Harald. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, Joost Zwarts, Irene Krämer & Gerlof Bouma (eds), 69–94. Amsterdam: Royal Netherlands Academy of Science.

Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage Publications.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Drach, Erich. 1937. *Grundgedanken der deutschen Satzlehre*. 4th edition. Frankfurt a. M.: Diesterweg.

Ellsiepen, Emilia & Bader, Markus. 2014. The under-additive effect of multiple constraint violations. *Cognitive Processing* 15(Suppl 1): S100–S102.

Fanselow, Gisbert. 2021. Acceptability, grammar,and processing. In *The Cambridge handbook of experimental syntax*, Grant Goodall (ed), 118–153. Cambridge: Cambridge University Press.

Featherston, Sam. 2021. Response methods in acceptability experiments. In *The Cambridge Handbook of Experimental Syntax*, Grant Goodall (ed), 39–61. Cambridge: Cambridge University Press.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, language, brain. Papers from the first Mind Articulation Project Symposium*, Alec Marantz, Yasushi Miyashita & Wayne O'Neil (eds), 95–126. Cambridge, MA: MIT Press.

Goodall, Grant. 2021. Sentence acceptability experiments: What, how, and why. In *The Cambridge handbook of experimental syntax*, edited by Grant Goodall, 7–38. Cambridge: Cambridge University Press.

Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.

Hawkins, John. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Heylen, Kris. 2005. A quantitative corpus study of German word order variation. In *Linguistic evidence. Empirical, theoretical and computational perspectives*, Marga Reis & Stephan Kepser (eds), 241–263. Berlin, New York: De Gruyter.

Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Hueber.

Keller, Frank. 2000. Gradience in grammar: Experimental and computational aspects of degrees of grammaticality. PhD dissertation. Edinburgh: University of Edinburgh.

Kempen, Gerard & Harbusch, Karin. 2005. The relationship between grammaticality ratings and corpus frequencies: A case study into word-order variability in the midfield of German clauses. In *Linguistic evidence. Empirical, theoretical and computational perspectives*, Marga Reis & Stephan Kepser (eds), 329–349. Berlin, New York: De Gruyter.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tubingen: Narr.

Lerot, Jacques. 1985. Zur Wortstellungsnorm im Deutschen. *Deutsche Sprache* 13(2): 137–142.

Müller, Gereon. 1999. Optimality, markedness, and word order in German. *Linguistics* 37: 777–818.

Olsson-Collentine, Anton, Van Assen, Marcel A. L. M. & Hartgerink, Chris H. J. 2019. The prevalence of marginally significant results in psychology over time. *Psychological science* 30(4): 576–586.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/ (20 February 2023).

Siewierska, Anna. 1993. On the interplay of factors in the determination of word order. In *Syntax: An international handbook of contemporary research* [Handbooks of Linguistics and Communication Science 9.1], Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds), 826–846. Berlin, New York: De Gruyter.

Sorace, Antonella. 2000. Gradients in auxiliary selection with intransitive verbs. *Language* 76: 859–890.

Stevens, Stanley S. 1957. On the psychophysical law. *The psychological review* 64: 153–181.

Vogel, Ralf. 2019. Grammatical taboos: An investigation on the impact of prescription in acceptability judgement experiments. *Zeitschrift für Sprachwissenschaft* 38(1): 37–79.

Wasow, Thomas. 2002. *Postverbal behavior*. Stanford: CSLI Publications.

Weskott, Thomas. 2021. Constituent order and acceptability. In *The Cambridge handbook of experimental syntax*, Grant Goodall (ed), 315–340. Cambridge: Cambridge University Press.

Weskott, Thomas & Fanselow, Gisbert. 2011. On the informativity of different measures of linguistic acceptability. *Language* 87(2): 249–273.