Chapter 8

On the (im-)possibility of reflexive binding into the subject of German experiencer-object verbs

Simon Masloch[a], Johanna M. Poppek[a] & Tibor Kiss[a]
[a]Ruhr-University Bochum

This paper presents an acceptability rating study on the possibility of reflexive binding into the subject of German experiencer-object psych verbs. Experiencer-object verbs are claimed to license exceptional binding patterns in many languages, but analyses differ in whether they relate this behaviour to a peculiar syntactic structure of the verbs or independently available logophoric binding. An explanation in terms of logophoricity is not viable in German, since the German reflexive *sich* does not allow a logophoric interpretation. The study shows that reflexive binding into the subject of German experiencer-object verbs is only possible in the midfield if the antecedent precedes the reflexive in linear order and – given the binary-branching structure of the midfield – thus c-commands it. The pattern observed poses a problem for predicate-based theories of binding and it is only explainable if sentence-level constituents in German are base-generated in their surface positions or scrambling does not reconstruct for binding.

1 Introduction

There is a long-standing debate about so-called "backward binding" into the subject of experiencer-object (EO) verbs, i.e. psych-verbs whose experiencer is realised as an object, (see i.a. Belletti & Rizzi 1988, Pollard & Sag 1992, Pesetsky 1995, Landau 2010, Cheung & Larson 2015). In Belletti & Rizzi's (1988) Italian examples in (1), only in the example containing an EO verb (1a) may the anaphor be bound although it is (superficially) not c-commanded by its antecedent.

Simon Masloch, Johanna M. Poppek & Tibor Kiss. 2025. On the (im-)possibility of reflexive binding into the subject of German experiencer-object verbs. In Gabriela Bîlbîie & Gerhard Schaden (eds.), *Empirical issues in syntax and semantics: Selected papers from CSSP 2023*, 197–222. Berlin: Language Science Press. DOI: 10.5281/zenodo.15450442

(1) Italian (Belletti & Rizzi 1988: 312)

a. Questi pettegolezzi su di sé preoccupano Gianni più di
these rumours about REFL worry Gianni more than
ogni altra cosa.
every other matter
'These gossips about himself worry Gianni more than anything else.'

b. * Questi pettegolezzi su di sé descrivono Gianni meglio di
these rumours about REFL describe Gianni better than
ogni biografia ufficiale.
every biography official
'These gossips about himself describe Gianni better than any official biography.'

Some authors take such examples to provide evidence for the unaccusativity of (certain classes of) EO verbs (e.g. Belletti & Rizzi 1988, Cheung & Larson 2015). Accordingly, the subject/nominative originates in a position below the object, such that c-command does hold at some point during the derivation. Others claim that such cases represent instances of logophoric or point-of-view-based binding, a phenomenon that extends beyond the domain of psych verbs (e.g. Pollard & Sag 1992, Bouchard 1995).

In this paper, we will present evidence from an acceptability judgment study that binding into the subject of EO verbs in the German midfield is possible only if the object precedes (and given the binary-branching structure of the midfield thus c-commands) the subject in surface order. In this regard, German is of special interest for multiple reasons: First, the overall grammaticality of examples analogous to (1a) is disputed (cf. Kiss 2012, Platzack 2012, Fischer 2015, Temme & Verhoeven 2017), with Fischer (2015) claiming that there is an effect of linear order. Secondly, despite the widespread assumption that scrambling disables binding possibilities in German (and enables new ones, see e.g. Haider 2017), Temme & Verhoeven (2017) claim to have found experimental evidence for backward binding with EO verbs in German using examples involving quantificational binding. Thirdly, German does not license logophoric binding (Kiss 2012), so if backward binding *is* possible, a logophoric interpretation of the reflexive cannot account for it. In the absence of an explanation relying on logophoricity, unaccusativity may be suggested to account for backward-binding patterns. However, what we find is that binding into the subject of an EO verb is possible in German only if it is *not* backward. The patterns observed can be explained by assuming that surface orders of the type A B imply that A asymmetrically

c-commands B and that the German reflexive *sich* requires a c-commanding antecedent. There is thus no need to return to unaccusativity, nor does an analysis suggest itself that is based on the concept of (lexical) predicates.

The structure of this paper is as follows: We will introduce some necessary background on German clause structure and linearisation as well as on binding peculiarities with EO verbs in Section 2. This will lead us to expectations about the acceptability of reflexive binding into the subjects of German EO verbs. Section 3 describes the experimental study, the results of which are discussed in Section 4. Section 5 concludes the paper.

2 Background

We will now briefly discuss some relevant aspects of German syntax and provide some background about "backward binding" with EO verbs.

2.1 German clause structure and the unmarked argument order with experiencer-object verbs

German is a verb-second language. The finite verb is placed after the first constituent in matrix clauses, but a verb-final order can be observed in embedded clauses. Placing a constituent in the pre-field (the area in front of the verb in verb-second clauses) may have interpretational effects (Frey 2006). Thus, in an experimental study all constituents of relevance should – if possible – be placed in the so-called midfield, i.e. the area between C (the position of the finite verb in verb-second clauses) and the verbal complex at the end of the clause.

Usually different linearisations of constituents in the midfield are grammatical, but there is a normal (information-structure-wise most neutral, see Höhle 1982/2019) order that is at least partially dependent on the predicate (we will use the terms *normal* and *unmarked* interchangeably). Deviations from the normal order outside a licensing context may influence acceptability judgments independently of binding constraints, making it necessary to consider their effects here. One prominent approach to German clause structure takes the unmarked order(s) to be base-generated while other orders are derived via scrambling (viewed as movement; see i.a. Frey 1993, Haider 2017). Other approaches favour base-generation of the different orders (i.a. Fanselow 2001) or assume a fixed base-generated order and movement, but do not equate it with the unmarked order (i.a. Müller 1999).

In the spirit of Belletti & Rizzi (1988), the literature on unmarked word order with EO verbs in German usually draws a distinction between those with an

accusative object and those with a dative object. Although the unmarked order with EO verbs is debated in the literature (cf. i.a. Lenerz 1977, Scheepers et al. 2000, Haider & Rosengren 2003, Ellsiepen & Bader 2018), recent experimental evidence points to a preference for object before subject with (most) dative-object EO verbs and a preference for subject before object with (most) accusative-object EO verbs if subjects are inanimate and all other factors potentially influencing linear order are controlled for (Temme & Verhoeven 2016, Masloch et al. 2024).

Masloch et al. (2024) assume base generation of sentence-level constituents and violable linear precedence constraints to account for their linearisation data. They treat linear precedence constraints as weighted constraints within Maximum Entropy Grammar (Goldwater & Johnson 2003), a probabilistic variant of Optimality Theory. There is much research on the factors influencing the linear order of elements in the German midfield (see i.a. Lenerz 1977, Uszkoreit 1987, Hoberg 1997, Keller 2000, Ellsiepen & Bader 2018). As Masloch et al. (2024) argue, most accusative-object EO verbs have a causer subject, while this is not the case for most dative-object EO verbs (their subject being an object of emotion in Pesetsky's 1995 terms). They do not assume constraints making reference to case or grammatical function, but (among others) a constraint CAUSER $\prec$ NON-CAUSER, which places causers before non-causers and has more weight than the constraint ANIMATE $\prec$ INANIMATE. Because in our experimental setting subjects will be inanimate but objects animate, these two constraints will lead to a preference for a subject before object linearisation with accusative-object EO verbs and a preference for object before subject with dative-object EO verbs. Since the constraints are violable, these preferences are not absolute and the reverse order is strictly speaking syntactically well-formed, although it may be less acceptable. We will follow this account. Furthermore, we follow Haider (2010) in assuming a binary-branching structure and the absence of functional projections between V and C. Thus, the schematic structure in Figure 1 emerges.

We will abstract away from certain phenomena, which are orthogonal to our analysis, such as the fronting of (reflexive) pronouns, the possibility to place constituents in the prefield, exceptional scope and binding options under a rise-fall intonation, and extraposition. Then, a phrase α dependent on a verbal head precedes another phrase β dependent on the same head in linear order iff α asymmetrically c-commands β. All orders of dependents of a verbal head are strictly speaking syntactically well-formed (unless they violate some other constraint, of course), but not all of them are equally acceptable in every context.

We will base the predictions for our experiment on this view of German clause structure. An approach assuming a fixed base order plus scrambling conceived

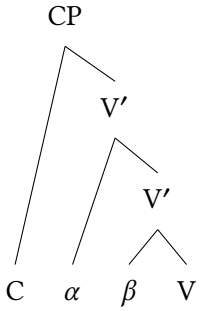

Figure 1: (Simplified) structural schema of German clauses

as movement to the left (as e.g. in Haider 2017) will produce the same predictions as long as binding constraints are evaluated at the target position only, i.e., i) scrambling does not reconstruct for reflexive binding; and ii) it is not the case that binding constraints can apply at any point.

2.2 Experiencer-object verbs and "backward binding"

It has been observed for many languages that an anaphor contained in the subject of an EO verb may precede its experiencer-object antecedent (see, among many others, for Italian Belletti & Rizzi 1988, for English Pesetsky 1995, for Chinese Cheung & Larson 2015). Examples like (1a), repeated here, pose a problem for theories of binding that require an anaphor to be c-commanded by its antecedent.

(1a) Italian (Belletti & Rizzi 1988: 312)
Questi pettegolezzi su di sé preoccupano Gianni più di ogni
these rumours about REFL worry Gianni more than every
altra cosa.
other matter
'These gossips about himself worry Gianni more than anything else.'

The solutions proposed in the literature can be divided into two broader classes: Those that take the backward binding pattern to relate to a peculiar syntactic structure of EO verbs (and possibly some wider class, e.g. Belletti & Rizzi 1988, Pesetsky 1995, Cheung & Larson 2015) and those that relate it to logophoric or point-of-view-based binding, which is available more generally (e.g. Pollard & Sag 1992, Bouchard 1995). On Belletti & Rizzi's (1988) account of the syntax of psych verbs, the nominative argument of EO verbs is not an external argument but originates in a position where it is c-commanded by the experiencer. It is then

assumed that Principle A can be satisfied before the stimulus moves to a position above the experiencer. While analyses that take the (surface) subject to originate in a position below the object are still widely assumed for dative-object EO verbs and sometimes for accusative-object EO verbs (see the overview in Rozwadowska et al. 2020), using backward binding to argue for it has somewhat fallen out of fashion (see, e.g. Landau 2010, Hirsch 2018). This is so because various authors have shown that backward binding can be licit even if it is impossible to establish a c-command relationship between the antecedent and the putative anaphor at any given syntactic level (see e.g. Pollard & Sag 1992, Bouchard 1995, Cançado & Franchi 1999). (2) is an illustrative example. Logophoric binding can account for such cases, so one needs to assume it anyway.

(2) Brazilian Portuguese (Cançado & Franchi 1999: 140)
Rumores sobre si explicam a insegurança mostrada por João.
rumors about himself explain the insecurity shown by John

Logophoric binding or exemption from Principle A, however, is not attested with the German reflexive *sich* (Kiss 2012). Picture-NPs do neither allow intersentential (3a), nor non-c-commanding (3b), nor split antecedents of embedded reflexives (3c), unlike in English (Pollard & Sag 1994: 245). The examples in (3) are not only somewhat degraded but grossly unacceptable on the coindexations given.

(3) (Kiss 2012: 158)

a. * Ulrich$_i$ war sauer. Ein Bild von sich$_i$ war beschädigt worden.
Ulrich was upset a picture of REFL had damaged been
'Ulrich was upset. A picture of himself had been damaged.'

b. * [Schumachers$_i$ Reklamevertrag] verlangte eine
Schumacher.GEN promotion.contract required a
Nacktaufnahme von sich$_i$.
nude.photo of REFL
'Schumacher's promotion contract required that nude photos of himself be taken.'

c. * Ulrich$_i$ zeigte Klaus$_j$ einige Bilder von sich$_{i+j}$.
Ulrich showed Klaus some pictures of REFL
'Ulrich showed Klaus some pictures of themselves.'

If backward binding with EO verbs is possible only due to exemption, it should thus *not* be possible in German. If it *is* possible, one could use this fact as an

argument for unaccusativity: On the assumption that scrambling reconstructs for binding or that a c-command requirement can be fulfilled at an early point or at any point in a derivation (as assumed by i.a. Grewendorf & Sabel 1999, Müller 1999), the experiencer could c-command the reflexive before the latter moves across it. If (dative) EO verbs have an unaccusative structure, binding into their subject should be possible irrespective of the linear order of the arguments. By contrast, only the orders in which the object precedes the subject containing the reflexive should be grammatical if there is a c-command requirement and all orders are base-generated as we assume here, or scrambling as movement destroys binding possibilities and creates new ones (as is frequently assumed, see e.g. Haider 2017).

The acceptability of German examples analogous to (1a) is disputed in the literature (cf. Kiss 2012, Platzack 2012, Fischer 2015, Temme & Verhoeven 2017). Fischer (2015) claims that there is an effect of linear order. According to her, binding into the subject of an EO verb is possible if the antecedent object-experiencer precedes it, and (4b) is acceptable to her.

(4) (Kiss 2012: 161, b. acceptable according to Fischer 2015: 1390)

a. * Ich glaube, dass die Bilder von sich den
I believe that the.NOM pictures.NOM of REFL the.DAT
Kindern gefielen.
children.DAT appealed.to

b. */✓ Ich glaube, dass den Kindern die Bilder von
I believe that the.DAT children.DAT the.NOM pictures.NOM of
sich gefielen.
REFL appealed.to
'I believe that the children liked the pictures of themselves.'

While Kiss (2012) judges (4b) as unacceptable, it is grammatical on his theory if his assumption that the subject must be the last argument to combine with the verbal projection is dropped and the linear order of constituents translates to c-command in the way we assume here (see Section 4.2).

The only experimental study on backward binding with EO verbs in German known to us is Temme & Verhoeven's (2017) and it claims that backward binding is more acceptable with EO than with action verbs. For reasons to be discussed below, the authors chose a configuration that does not involve reflexive, but quantificational binding. They report the results of two experiments (one for accusative-object verbs, one for dative-object verbs), in which they compared

the acceptability of backward binding into the subjects of EO verbs and agentive verbs in two conditions: particular and generic. Since they argue that apparent binding possibilities on the latter reading are only illusory, we will focus on the former. Participants were asked to provide binary acceptability judgments. In both experiments, Temme and Verhoeven found a significant (and non-negligible) effect of verb class to the extent that backward binding was more acceptable with EO verbs. However, the overall acceptability within the EO conditions was still not high (30 % for accusative-object, 40 % for dative-object EO verbs, which compare to an average acceptability of around 20 % for Principle-C violations and ca. 83 % for backward coreference across their experiments). Temme & Verhoeven (2017) rightfully argue that it is the observed difference between the conditions in the controlled experiment that counts and that the relatively low acceptance rate does not imply ungrammaticality. They propose that it may be due to processing difficulties that arise with quantificational binding and the backward dependency as well as the fact that the reading of the stimuli they asked their participants to evaluate is not the most prominent one. Ultimately, they take their findings to show that backward binding is a peculiar property of psych verbs in German after all. Yet, we take it to be possible that other factors are responsible for the effect. In particular, Webelhuth (2022) recently showed in a corpus study that quantificational binding in German is possible *without* c-command. He concludes that "[t]he overall picture that emerges from the corpus evidence is thus that topicality motivates wide scope and scope rather than c-command licenses […] bound pronouns" (Webelhuth 2022: 387). In an article about argument linearisation, Temme & Verhoeven (2016) argue that experiencers are more likely to be aboutness-topics than patients. Thus, it may be the case that Temme & Verhoeven's (2017) results are due to the experiencer being more likely to be interpreted as the topic than a patient and taking wide scope in turn, which licenses quantificational binding. Since Webelhuth (2022) does not claim reflexive binding to be licensed by topicality, we consider it preferable to rely on reflexive binding to test for the availability of backward binding in German.

To sum up: The data on the possibility of backward binding in German is murky. The acceptability of pertinent examples is disputed in the literature and the only experimental study finds an effect, but it is weaker than one may expect and may be caused by an independent factor. Logophoricity is not a factor in German. The possibility of binding into a subject preceding the object could be explained by assuming unaccusativity and a c-command condition that either allows reconstruction or may be fulfilled at any or an early point. Both examples in (4) should be grammatical then. If arguments are base-generated in their sur-

face positions or scrambling does not reconstruct for binding, only (4b) should be grammatical. Masloch et al. (2024) argue that the latter view explains the linearisation preferences they observe with German EO verbs less naturally.

A potential problem for an experimental investigation into the possibility of reflexive backward binding in German pointed out by Temme & Verhoeven (2017: 286) concerns the subjects themselves: Since German lacks a genitive reflexive, a reflexive can only be embedded in the subject within a PP. However, the usage of such a PP can be functionally overshadowed by a considerably more frequent construction involving a possessive, as in (5).

(5) (based on an in-text example by Temme & Verhoeven 2017: 286)
Er betrachtete seine Möbel / ??die Möbel von sich.
he beheld his furniture the furniture of REFL
'He looked at his furniture.'

Such overshadowing may lead to reduced acceptability of the stimuli in an experimental setting and should thus be avoided. However, not all [N [P REFL]] structures are equally unacceptable: e.g. *Bilder von sich* 'pictures of REFL' as in (4) is not generally unacceptable as shown by sentences like *Warum hat Claude Cahun$_i$ die Bilder von sich$_i$ zurückgehalten?* 'Why has Claude Cahun withheld the pictures of herself?' (Kiss 2012: 156). In this sentence, *ihre Bilder* 'her pictures' would mean something like 'pictures she made/owns' rather than 'pictures of herself / pictures depicting herself': It is one of the cases often noted in the literature on possessives (see e.g. Barker 2019) where the exact relation holding between the possessor and the possessed is provided by context (Claude could e.g. own the pictures or it could be pictures she has taken). *Bilder* 'pictures' being a relational noun, the question arises why the interpretation 'pictures depicting her' is not salient (or perhaps unavailable) for *ihre Bilder*. We will tentatively assume that the preposition *von* 'of' is not devoid of meaning and that NPs containing such PPs compete with NPs containing a possessive. Which variant is preferred for a given meaning will then depend on the exact literal meaning of the candidates, the factors involved in the competition and general pragmatic principles. For the purposes of experiments on reflexive binding, the exact analysis of this phenomenon does not matter as long as there is a way to ensure that the [N [P REFL]] structures are not overshadowed by [POSS N] structures (but see Section 4).

3 Experimental Study

We aimed to answer the question if reflexive binding into the subject of EO verbs is possible in German by conducting an acceptability rating study. The study has been preregistered with OSF (https://doi.org/10.17605/OSF.IO/EV7MA). All scripts and materials are available via https://doi.org/10.17605/OSF.IO/VNWFQ.

3.1 Design

The design reflects the two factors ORDER (*subject before object*, *SO* or *object before subject*, *OS*) and CASE (of the object, *accusative* or *dative*). While CASE is tested between items (since there is no synchronic object-case alternation with EO verbs having a subject in German), each item is presented in both ordering conditions. Participants only see each item in one ordering condition, but each of them rates the same number of *SO* and *OS* sentences. Answers are provided on a 5-point scale ranging from *vollkommen unnatürlich* 'completely unnatural' to *vollkommen natürlich* 'completely natural'. All points had a natural language name.

3.2 Materials

Test items were constructed according to examples (6–7), presented here without any acceptability judgment:

(6) Accusative object

a. Subject before object:
Es ist offensichtlich, dass das Gerücht über sich den
it is obvious that the.NOM rumour.NOM about REFL the.ACC
Professor genervt hat.
professor.ACC annoyed has

b. Object before subject:
Es ist offensichtlich, dass den Professor das Gerücht
it is obvious that the.ACC professor.ACC the.NOM rumour.NOM
über sich genervt hat.
about REFL annoyed has
'It is obvious that the rumour about himself annoyed the professor.'

(7) Dative object

a. Subject before object:
Es ist allgemein bekannt, dass die Meldung über sich
it is commonly known that the.NOM report.NOM about REFL
dem Opernsänger gefallen hat.
the.DAT opera.singer.DAT appealed.to has

b. Object before subject:
Es ist allgemein bekannt, dass dem Opernsänger die
it is commonly known that the.DAT opera.singer.DAT the.NOM
Meldung über sich gefallen hat.
report.NOM about REFL appealed.to has
'It is commonly known that the opera singer liked the report about himself.'

Test items contained the clause of interest embedded in a matrix clause to ensure a verb-final sentence. In total, we used eight test items containing an accusative-object EO verb and eight test items containing a dative-object EO verb, each of them in both ORDER variants.[1] We created two lists so that each participant rated only one ordering variant per item. In addition to the test items, the questionnaires contained 64 filler items, so that each participant judged 80 sentences (plus one sentence in the instructions). Among the fillers, there were six unannounced calibration items included to familiarise participants with the task and the scale, sixteen control items used to identify uncooperative or distracted participants, and five attention items used to detect inattentive participants. Filler items varied in expected acceptability so that participants would see roughly the same number of clearly acceptable, clearly unacceptable, and somewhat degraded sentences. Within each subcategory, half of the filler items were related to the test items either by containing *sich*, by containing a noun-preposition-noun structure, or by containing a psych verb, while the other half was unrelated. All items were presented in pseudo-randomised order subject to some constraints.[2]

[1]A complete list of all items used in the study can be found in the OSF directory. Dative-object verbs: *auffallen* 'to strike', *behagen* 'to please', *einleuchten* 'to be evident', *gefallen* 'to appeal to', *imponieren* 'to impress', *missfallen* 'to displease', *nahegehen* 'to afflict', *widerstreben* 'to have an aversion against'; accusative-object verbs: *anekeln* 'to sicken', *ärgern* 'to anger', *ängstigen* 'to frighten', *beeindrucken* 'to impress', *befremden* 'to alienate', *faszinieren* 'to fascinate', *nerven* 'to bother', *verärgern* 'to annoy'.

[2]The presentation of items was constrained as follows: calibration items had to come first, test and control item had to be separated by fillers, controls had to occur in the last 66 % of the questionnaire, and the last item had to be a filler.

The verbs used in the test items were chosen based on their syntactic behaviour in corpus data – essentially following the procedure of and using the materials from Masloch et al. (2024) – so that a preference for inanimate subjects, the frequency of non-psych readings and other potential confounding factors were taken into account. In all test items, the subject was an NP containing an embedded PP whose internal argument was the third person reflexive *sich*, while the embedded verbs' object was the only possible antecedent for the reflexive. The noun-preposition sequences are frequent collocates and we ensured that the use of the PP is not overshadowed by a possessive construction (as in 5). In order to do so, nouns that frequently have a preposition as their right neighbour were extracted from DeReKo (Kupietz et al. 2010) using KorAP (Diewald et al. 2016). From these, 327 nouns were manually chosen. We then calculated collocation scores with 81 prepositions (as direct right neighbours) and possessive pronouns (maximally three words to the left of the noun) for each of them and chose noun-preposition combinations from the pairs with a high logDice (an association score defined by Rychlý 2008). Afterwards we manually checked whether the use of an embedded reflexive is overshadowed by a possessive construction (cf. Section 2.2). The noun-preposition-reflexive combinations used in the items were chosen such that there is no overshadowing in our judgment. Additionally we avoided psych nouns for potential confounding effects.

3.3 Participants and Procedure

Participants (monolingual native speakers of German, residents of Germany, Austria, or Switzerland) were recruited via *Prolific* (prolific.com). 79 participants completed the questionnaire and received a compensation of £ 3.5. A typical run lasted ca. 15 minutes. The experiment was conducted using a web-based infrastructure using jsPsych (de Leeuw 2015) and JATOS (Lange et al. 2015) on a university server – where participants' individual reaction times were automatically measured. Taking the control- and attention checks specified in the preregistration as well as possible topic awareness (checked with an open question at the end of the survey) into account, data from 48 participants was included in the analysis.[3] After giving their informed consent to participate in the experiment, participants read written instructions asking them to rate how natural the

[3] A version of the analysis script where all participants are included is available via the OSF directory and the results are interpretation-wise the same as the ones presented here. The reason for the comparatively large number of participants whose data did not enter the analysis were the rather strict predefined exclusion criteria, namely: 1. the participant guessed the topic of the study correctly or displayed significant linguistic knowledge (we asked participants to guess the topic), 2. the participant did not complete the questionnaire, 3. the participant

sentences sound to them as sentences of their mother tongue. They saw an example item on the instructions page. The experiment started with the six unannounced calibration items. Each item was presented on its own page together with the answer options. There was no time limit for providing an answer.

3.4 Hypotheses and predictions

As discussed in Section 2.1, we follow Masloch et al.'s (2024) account of argument linearisation in the midfield for German EO verbs, which takes surface-order at face value: A constituent α in the midfield is taken to c-command a constituent β if and only if α precedes β in linear order. Combined with the assumption that the German reflexive *sich* must be c-commanded by its antecedent, it follows that sentences in which the subject precedes the object are ungrammatical (because the reflexive cannot be c-commanded by its antecedent).[4]

(8) Main hypothesis
In the German midfield, the object of an experiencer-object verb cannot bind a reflexive embedded in a subject preceding it.

Thus, an item is ungrammatical if the subject precedes the object in the target clause. As mentioned in Section 2.2, theories of German clausal syntax that take scrambling not to reconstruct for reflexive binding will share this hypothesis irrespective of the base structure assumed. Sentences in which the object precedes the subject are strictly speaking grammatical but may violate linear precedence constraints, possibly leading to different degrees of acceptability. An OS linearisation will violate CAUSER $\prec$ NON-CAUSER with (most) accusative-, but

did not judge at least 80 % of the attention items correctly, 4. the participant did not judge at least 80 % of the related control items correctly, 5. the participant did not judge at least 80 % of the unrelated control items correctly, 6. the participant had unusually long or short answering times as determined by Pieper et al.'s (2025) method, 7. the participant self-reported residing in a country or area where German is not the official language. The OSF directory contains the script that was used in the exclusion process, where all exclusions are discussed. We slightly deviated from Pieper et al.'s (2025) criteria for reaction times, which appeared to be too strict given the overall very fast reaction times. Overall, these criteria are quite strict because participants have to fulfill all of them. We think that this is a desirable property because participants recruited via web-based participant recruitment platforms tend to rush through studies and the decisive manipulation was rather small and could easily be overlooked.

[4] As mentioned in Section 2.1, we do not aim at covering pronoun fronting, which is a different mechanism than the word order freedom we look at here (Haider 2017). We assume that *sich* is fronted in many apparent counterexamples to a c-command condition. For our purposes, the precise mechanism behind such cases is irrelevant since in our items the reflexive is embedded.

Table 1: Predicted acceptability of the test items within the different conditions (see main text for qualification)

	dative	accusative
OS	high	medium
SO	low/medium	low

not with (most) dative-object EO verbs (see Section 2.1). An SO linearisation violates ANIMATE ≺ INANIMATE, which is outweighed by CAUSER ≺ NON-CAUSER with accusative-object verbs, but not with dative-object verbs. Thus, without any violations of binding constraints OS should be more acceptable with dative-object verbs, SO with accusative-object verbs.

Based on these prerequisites, we expect dative-object verbs to receive high ratings in OS linearisation while the same order is marked with accusative-object verbs, which should result in lower ratings. In SO order, the reflexive is not c-commanded by its antecedent and the order is marked for dative-object verbs. However, we may hypothesise that some participants, especially when they try to behave like cooperative discourse-participants (remember that we asked participants to rate how natural the sentences sound to them) or parse the sentences only superficially, will correct the ungrammatical sentence. Since the resulting OS order is unmarked with dative EO verbs, participants may rate the items containing them higher than one would expect based on their theoretical grammaticality status. In contrast, since SO already is the unmarked order with accusative-object verbs, this effect is not possible for them. Consequently, sentences in this condition should be rated as unnatural. This state of affairs is summarised in Table 1.

Given the model we will use for the analysis (see Section 3.5), our assumptions lead to the following expected effects:[5]

(9) Expectations fixed effects:
- CASE: Mildly positive
- ORDER: Medium/strong positive
- CASE × ORDER: marginal or non-existent

[5]This is *not* the model mentioned in the pre-registration. A reviewer for CSSP suggested that a sum-coded model may be easier to understand than the dummy-coded one we were using originally. We think that they are right and only discuss the sum-coded model here. The dummy-coded model is still available in the analysis script on the OSF directory.

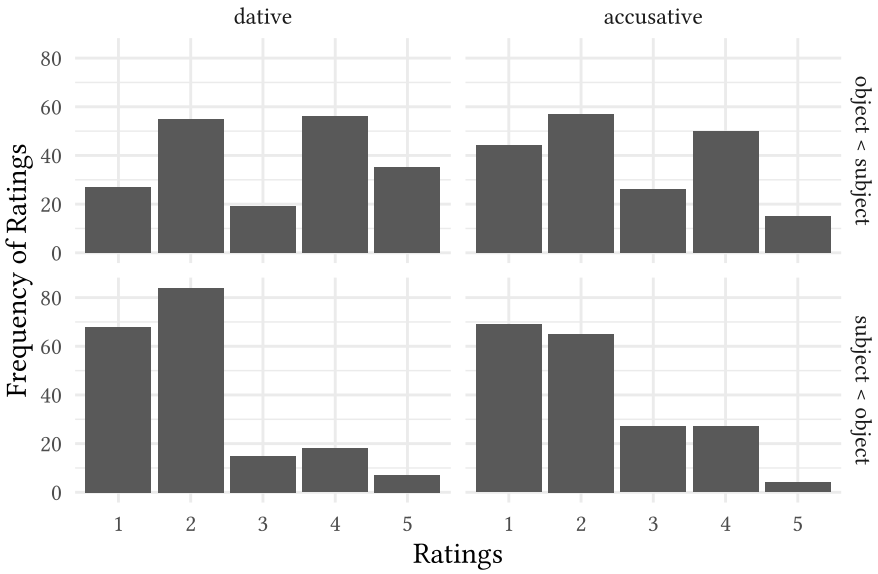

Figure 2: Empirical distribution of ratings. "5" stands for "completely natural", "1" for "completely unnatural". All choice options were presented to the participants with a natural language label.

3.5 Results

Figure 2 displays the empirical distribution of ratings in all four conditions. We see that sentences in which the subject containing the reflexive precedes the object (condition *SO*) received very low ratings, although ratings (unexpectedly) improve slightly with accusative-object verbs. In the *OS* condition, in which the reflexive is preceded by its antecedent, sentences receive overall better judgments, although there is still a large number of lower ratings. Ratings are higher with dative-object verbs there. These descriptive results speak in favour of the main hypothesis (8).

To model the data and test our hypotheses, we fitted a Bayesian cumulative logit generalised linear mixed model with flexible thresholds using the `brms` package (Bürkner 2017) in R (R Core Team 2023). Cumulative models are a type of model that can be used to analyse ordinal data (see Bürkner & Vuorre 2019 for an introduction): Since one cannot assume that the intervals between response options of Likert items have equal size, it is not appropriate to use a metric model (see i.a. Liddell & Kruschke 2018). Both factors were sum-coded with *dative* and *OS* coded 1 and *accusative* and *SO* coded −1. The model includes fixed effects

for CASE, ORDER, and their interaction, varying intercepts for participants and items, a varying slope for ORDER for items, and varying slopes for CASE, ORDER, and their interaction for participants, as well as all possible correlation parameters between them.[6] In Bayesian models, parameters are random variables, so one can talk about the credibility of different values (an introduction aimed at a linguistic readership is provided by Nicenboim et al. 2024). One starts off with a prior distribution across the parameters, which reflects one's prior knowledge, and updates it using the data to receive a posterior distribution, which reflects the uncertainty about the parameter values. Given the lack of previous comparable studies and quantitative predictions for the effect sizes, we use mildly informative, theory-neutral regularising priors.[7] For hypothesis testing, we will use Bayes factors: The Bayes factor (BF) is the ratio of the marginal likelihoods of two models and tells us under which of them the data are more likely. We will compare the model we report here against a model in which the effect of interest is set to 0. According to Jeffreys (1939, cited after Nicenboim et al. 2024), a BF of 3 indicates moderate evidence in favour of the first model, a BF of 10 strong evidence and a BF of 100 extreme evidence. 1 is the neutral value, ⅓ indicates moderate evidence in favour of the second model etc.[8]

There is an effect of ORDER in the expected direction ($\hat{\beta} = 0.79$, 95 % credible interval (CrI) = $[0.47, 1.12]$)[9], which is, however, not quite as strong as expected as it roughly corresponds to the difference between two levels of the ordered response variable.[10] Nevertheless, the Bayes factor shows that there is extreme evidence for this effect (Bayes factor computed as Savage-Dickey density ratio

[6]The analysis script on the OSF directory contains several additional models as well as a comparison between them. We will focus on this model here.

[7]We performed prior predictive checks and additionally fitted models with different (and more informative) priors, which can be found in the analysis script on the OSF directory. The results are interpretation-wise the same.

[8]The Bayes factor depends on the prior. The analysis script contains a sensitivity analysis for the Bayes factors as well as a detailed look at the posterior distributions of different models (including models with more informative priors). For our theoretical purposes, the results are similar.

[9]We use hats (ˆ) to indicate estimates. When describing the marginal posterior distribution of a parameter in Bayesian models, we use the mean as the measure of central tendency (estimate) and the standard deviation for variability (CrIs).

[10]Due to the sum-coding used, the distance between *OS* and *SO* would be $0.79 \times 2 = 1.58$, which is smaller than the distance between the first and the second or between the third and the fourth threshold, but larger than the distance between the second and the third on the latent variable within the model, see the full model summary in the analysis script on the OSF. In a cumulative model, the response is taken to relate to a latent variable – in our case: perceived naturalness – that can be modelled as linear and is partitioned into ordered bins corresponding to the response options via thresholds that are estimated in the model.

between the model and a model where β is set to 0; $BF_{10} = 126.7$). Changing from *SO* to *OS* leads to an increased naturalness. The point estimate for the effect of CASE ($\hat{\beta} = 0.17$, CrI = $[-0.24, 0.57]$, $BF_{10} = 0.074$) is slightly positive, but we do not have enough evidence to postulate its existence. Indeed, the Bayes factor shows that one may do better without it. In any case, the effect is only small. The reason we expected a mildly positive effect of CASE was that OS is unmarked for dative-object EO verbs, but not for the accusative-object ones. While descriptively sentences are rated higher in *dative OS* than in *accusative OS*, they are rated lower in *dative SO* than in *accusative SO*. Judments of naturalness may correspond to normal order in a more straightforward way, such that a sentence perceives higher ratings irrespective of binding constraints if it is normally ordered (this may be a caused by performance mistakes). The interaction effect would capture such a pattern: a positive value of it would correspond to a preference for the normal order irrespective of the other factors including binding constraints because – given the encoding chosen – *dative SO* and *accusative OS* get the value -1 and *dative OS* and *accusative SO* get the value 1. Although the posterior distribution hints at a small positive effect ($\hat{\beta} = 0.23$, CrI = $[-0.06, 0.52]$), the Bayes factor indicates that the data provide evidence against it ($BF_{10} = 0.138$).

The model assumes a comparatively large standard deviation of the participants' varying intercepts ($\hat{SD} = 1.57$, CrI = $[1.22, 2]$), so there is variability between participants. The standard deviation of the participants varying slope for ORDER is considerable ($\hat{SD} = 0.5$, CrI = $[0.26, 0.75]$), while the ones for CASE ($\hat{SD} = 0.16$, CrI = $[0.01, 0.39]$) and for the interaction ($\hat{SD} = 0.22$, CrI = $[0.01, 0.48]$) are not. There is variation between items ($\hat{SD} = 0.74$, CrI = $[0.45, 1.18]$), also for the varying slope for ORDER ($\hat{SD} = 0.47$, CrI = $[0.23, 0.81]$). The estimated correlations between varying effects are rather unremarkable.

An exploratory look into the responses of individual participants shows that twelve out of the 48 participants whose responses entered the analysis assign low scores across conditions. Additionally, three of the test items received almost only low scores in both ordering conditions. Taken together, these observations may explain the overall lower level of acceptability (and hence the surprisingly small effect size of ORDER) and in part also the variation among participants and items. If the relevant items and participants are excluded, accusative-object EO verbs receive mixed judgments in the *OS* condition, dative-object EO verbs rather good ones.

4 Discussion

4.1 General discussion

We take these results to support our main prediction: Reflexive binding into the subject of German EO verbs is licit only if it is *not* backward.[11] This holds for dative-object as well as for accusative-object EO verbs. As discussed in Section 2.2, this is expected on different accounts of German clausal syntax and the syntax of EO verbs, but it is *not* expected on an account that i) takes (dative) EO verbs to be unaccusative (which should translate into an OS base order); and ii) takes scrambling to reconstruct for reflexive binding. Indeed, the results are incompatible with the idea that scrambling reconstructs for binding because if it did so, there should be no difference between the ordering conditions. Thus, our data do not speak against the unaccusativity hypothesis (taken to imply an OS base order) directly because the base order does not matter for reflexive binding if scrambling does not reconstruct for it; but they also do not speak for it. Both a base-generation approach as described in Section 2.1 and a scrambling-as-movement account without reconstruction for reflexive binding are compatible with the data.

We expected a mildly positive effect of CASE, which seems to be inexistent. The reason we expected this effect was that we assumed that participants may rate a sentence better if the binding condition is fulfilled in the unmarked order. This does not seem to be the case.

Two remarkable aspects are the unexpectedly low level of acceptability (also in comparison to filler and control items categorised as acceptable a priori, which received high ratings, see the analysis script on the OSF directory) and the substantial individual variation. We do not have a full explanation for this at the moment, but regarding both, one has to consider that the test items were complex sentences that had to fulfill highly specific criteria and contained a relatively infrequent phenomenon, namely the PPs containing reflexives discussed in Section 2.2. The overshadowing process mentioned there presumably involves competition between a PP containing the reflexive and the possessive, but a PP containing a personal pronoun will be involved as well. We suspect that there are

[11] For a discussion of the surprisingly low overall acceptability in the OS condition, see below. One may object that both orderings could be ungrammatical and that enhanced acceptability of the OS order is due to priming effects: Since participants see the experiencer object first, they are already familiar with the possible antecedent and will be able to correct the grammar violations instantly, whereas this is not the case in the SO variant. An account along such lines is not tenable since the examples in (3), where the possible antecedent precedes the reflexive, too, are clearly unacceptable.

differences both between and within speakers (varieties, register) with respect to the weighting of the factors involved. These will influence how natural the noun-preposition-reflexive combinations sound to them in the given setting because (as one may assume) the losing candidate in a competition will be judged as less natural. Items may differ in which factors involved in the competition are relevant. To the best of our knowledge, there is no in-depth investigation of the phenomenon yet.[12] The fact that the standard deviation of the participants' varying slope for ORDER is not as low as one may expect may be due to the participants who rate all test items as unnatural. Something similar may happen with the standard deviation of the items' varying slope for ORDER since there are some test items that received low ratings across conditions. A radical alternative explanation for the rather low overall acceptability and the individual variation would be lectal variation in binding domains such that for some speakers the reflexive has to be bound only at clause level while for others the binding domain is narrower, so that the reflexive needs an antecedent within the NP or PP. However, given the availability of a plausible alternative, we do not want to pursue this path.

A reviewer asks us to disentangle the effects of linear order, c-command and topicality, which all seem to make the same prediction here. While we only looked at the midfield, where linear order and c-command (largely) correspond to each other, a constituent situated in the prefield linearly precedes the midfield. It is widely assumed that the prefield constituent moves there from the midfield. It may thus (at some point) be c-commanded by a constituent that follows it in surface order. (10) is acceptable even though the reflexive precedes its antecedent.

(10) [Bilder von sich$_i$ veröffentlicht]$_1$ hat [Claude$_i$ [fast nie t_1]].
pictures of REFL published has Claude almost never
'Claude almost never published pictures of herself.'

In (11), the reflexive in the accusative object may be bound by the subject or the dative object, although only the subject is topical.

(11) *Was ist mit der Fotografin?* 'What about the photographer?'
Die Fotografin$_i$ zeigte einem Kunden$_j$
the.NOM female.photographer.NOM showed a.DAT customer.DAT

[12] The EISS reviewers asked for baselines involving PPs containing full NPs or personal pronouns. We did not include items containing such NPs to keep the number of conditions low and because we did not expect the overall rather low acceptability. On our explanation of the latter, such items could not function as real baselines (because they either stand in competition with the variant we used or lack a competing candidate) but would be very interesting for further investigation.

Bilder von sich$_{i/j}$.
pictures.ACC of REFL
'The photographer showed a customer pictures of herself/himself.'

Both linear order and topicality are thus unlikely to be the decisive factor for licensing reflexive binding by themselves.

4.2 Theories of binding

So far, we assumed that the German reflexive *sich* has to be c-commanded by its antecedent, which aligns with Principle A of classical binding theory. Theories that try to capture binding data using tree-configurational notions such as c-command are rivaled by predicate-based theories of binding (i.a. Pollard & Sag 1992, Reinhart & Reuland 1993), in which co-argumenthood is decisive. On Reinhart & Reuland's (1993) account, only heads with an external argument count as syntactic predicates. According to their condition A, reflexive syntactic predicates (i.e., predicates that have two co-indexed arguments) need to be reflexive marked, which can either happen lexically or via a SELF-anaphor. German *sich* can be a SELF-anaphor according to Reuland & Reinhart (1995). Pollard & Sag (1992) define binding conditions in terms of relative obliqueness. In their analysis of English anaphora, an anaphor has to be co-indexed with a less oblique co-argument if there is one. In our test items, the predicate relevant to determining co-argumenthood is the noun or the preposition. In both cases, there is no co-argument/subject. Thus, it is not a syntactic predicate for Reinhart & Reuland (1993) and condition A does not apply. Thus, the anaphor does not have a less oblique co-argument so that Pollard & Sag's (1992) Principle A does not apply. As a result, the reflexive should be licensed without being bound.[13] Even if there were some further principle requiring *sich* to have a co-argument or if one were to assume that there is an unpronounced external argument of the noun or preposition, one could not capture our data, since it is the positioning of the whole syntactic predicate (the NP/PP) that makes the difference. Hence some structural condition must be at play.[14]

[13] Pollard & Sag (1992) do not claim their theory to be applicable to languages other than English. Thus, our data do not speak against their theory directly, they only show that it cannot trivially be extended to German.

[14] According to Reuland & Reinhart (1995) and Reuland (2011), *sich* may also be a SE-anaphor, in which case it would not mark the predicate as reflexive. However, it should not be possible to stress SE-*sich* (see Reuland & Reinhart 1995: 249 sqq., Reuland 2011: 275 sqq.), but in our judgment stressed *sich* is perfectly fine in examples like (4b). In order to be interpreted as bound, SE-*sich* would have to be in a chain with its antecedent. This cannot be the case (if only

This structural condition need not be a universal principle: On Kiss' (2012) account of reflexive binding, anaphoric dependencies are introduced in syntax. A reflexive pronoun will receive a feature **D**(*n*), where *n* is an index, if it is an argument of a head with an articulated argument structure (≈ having an external argument, see Kiss 2012), which will bear the feature +ARG-S. **D**(*n*) is projected upwards until *n* is identified with the index of a sister of a phrase bearing the feature. In German, a feature **Đ**(*n*) representing an inactive dependency is introduced if the head does not have an articulated argument structure (−ARG-S), and projected in the same way until it is activated (= leads to the phrase having the feature **D**(*n*)) when meeting a +ARG-S head.[15] A local resolution condition requires dependencies to be resolved within the clause. In effect, this means that German *sich* has to be co-indexed with a c-commanding phrase within the same clause.[16] Figure 3 illustrates how this works for (4b). Since P and N do not have an articulated argument structure there, the **Đ**(*n*) introduced with *sich* becomes active only once $_2$NPbecomes a daughter of the verbal projection, i.e. $_2$NP bears **D**(*n*). **D**(*n*) is projected upwards to $_1$V′, where *n* can then be identified with the index of $_1$NP. Indeed, it must do so in order for the local resolution condition to be fulfilled. By contrast, local resolution cannot be fulfilled if the NP containing the reflexive is the last one to combine with the verbal projection as in (4a) and our SO items, leading to ungrammaticality.

5 Summary

Our study shows that reflexive binding into the subject of experiencer-object verbs is licensed in the German midfield only if the subject is preceded – and thus c-commanded – by the antecedent in surface structure. The results are in principle compatible with both free base-generation and movement-based accounts of linearisation in the midfield, but with the latter only if scrambling is taken not to reconstruct for binding. Analysing German EO verbs as unaccusative is not necessary to explain their reflexive binding patterns (although they are not

because Reuland (2011: 167 sqq.) takes D to block the attraction process that would be necessary on his account). If SE-*sich* cannot enter a chain, one may expect a logophoric interpretation to occur, but 1. German *sich* does not have a logophoric interpretation as shown in Section 2.2 and 2. one would get the same problems as mentioned for SELF-*sich* in the main text then.

[15] English is taken to lack inactive dependencies, so something like a predicate-based binding theory emerges.

[16] Note that on Kiss' (2012) theory there may also be differences between different lexical items within a language, so our results for *sich* may not be directly transferable to reciprocal *einander* 'each other'.

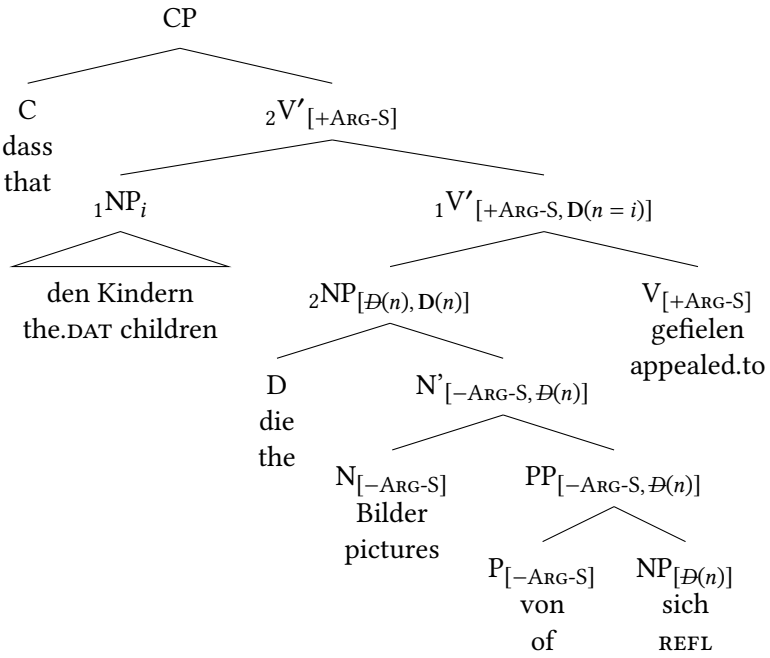

Figure 3: Analysis for (4b). Numbers are used to distinguish nodes in the tree, small letters for indices

incompatible with unaccusativity). Because the positioning of the constituent containing the embedded reflexive influences acceptability, the results are problematic for predicate-based binding theories.

Acknowledgements

We would like to thank three anonymous reviewers for CSSP and three anonymous reviewers for EISS as well as the audience at CSSP and the editors of this volume, Gabriela Bîlbîie and Gerhard Schaden, for their valuable feedback and suggestions! A special thanks goes to Tamara Stojanovska for performing the collocation analyses needed for item construction. The research reported here was supported by the Deutsche Forschungsgemeinschaft (DFG) grant no. 437144413 to Tibor Kiss.

Abbreviations

BF	Bayes factor	SO	subject before object
CrI	95 % credible interval	OS	object before subject
EO	experiencer-object		

References

Barker, Chris. 2019. Possessives and relational nouns. In Paul Portner, Klaus von Heusinger & Claudia Maienborn (eds.), *Semantics: Noun phrases and verb phrases* (Handbooks to Semantics 4), 177–203. Berlin, Boston: de Gruyter.

Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language & Linguistic Theory* 6(3). 291–352. DOI: 10.1007/BF00133902.

Bouchard, Denis. 1995. *The semantics of syntax: A minimalist approach to grammar*. Chicago, IL: The University of Chicago Press.

Bürkner, Paul-Christian. 2017. Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). DOI: 10.18637/jss.v080.i01.

Bürkner, Paul-Christian & Matti Vuorre. 2019. Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science* 2(1). 77–101. DOI: 10.1177/2515245918823199.

Cançado, Márcia & Carlos Franchi. 1999. Exceptional binding with psych-verbs? *Linguistic Inquiry* 30. 133–143. DOI: 10.1162/002438999553995.

Cheung, Candice Chi-Hang & Richard K. Larson. 2015. Psych verbs in English and Mandarin. *Natural Language & Linguistic Theory* 33(1). 127–189. DOI: 10.1007/s11049-014-9259-3.

de Leeuw, Joshua R. 2015. jsPsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods* 47(1). 1–12. DOI: 10.3758/s13428-014-0458-y.

Diewald, Nils, Michael Hanl, Eliza Margaretha, Joachim Bingel, Marc Kupietz, Piotr Bański & Andreas Witt. 2016. KorAP architecture: Diving in the deep sea of corpus data. In *Proceedings of the 10th international conference on language resources and evaluation (LREC 2016)*, 3586–3591. Portorož/Paris: European Language Resources Association (ELRA).

Ellsiepen, Emilia & Markus Bader. 2018. Constraints on argument linearization in German. *Glossa* 3(1). 1–36. DOI: 10.5334/gjgl.258.

Fanselow, Gisbert. 2001. Features, θ-roles, and free constituent order. *Linguistic Inquiry* 32(3). 405–437. DOI: 10.1162/002438901750372513.

Fischer, Silke. 2015. Theories of binding. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science [HSK] 42), 1357–1400. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110363708-016.

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation* (Studia Grammatica 35). Berlin: Akademie Verlag.

Frey, Werner. 2006. Contrast and movement to the German prefield. In Valéria Molnár & Susanne Winkler (eds.), *The architecture of Focus* (Studies in Generative Grammar 82), 235–264. Berlin, New York: Mouton de Gruyter.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm workshop on 'Variation within Optimality Theory'*, 113–122. Stockholm.

Grewendorf, Günther & Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language & Linguistic Theory* 17. 1–65. DOI: 10.1023/A:1006068326583.

Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511845314.

Haider, Hubert. 2017. Mittelfeld phenomena: Scrambling in Germanic. In *The Wiley Blackwell companion to syntax*, 2nd edn., 2573–2645. Wiley. DOI: 10.1002/9781118358733.wbsyncom048.

Haider, Hubert & Inger Rosengren. 2003. Scrambling: Nontriggered chain formation in OV languages. *Journal of Germanic Linguistics* 15(3). 203–267. DOI: 10.1017/S1470542703000291.

Hirsch, Nils. 2018. *German psych verbs: Insights from a decompositional perspective*. Humboldt-Universität zu Berlin. (Doctoral dissertation).

Hoberg, Ursula. 1997. Die Linearstruktur des Satzes. In Gisela Zifonun, Ludger Hoffmann & Bruno Strecker (eds.), *Grammatik der deutschen Sprache*, 1495–1680. Berlin: De Gruyter.

Höhle, Tilman N. 2019. Explikationen für „normale Betonung" und „normale Wortstellung". In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur deutschen Grammatik: Gesammelte Schriften von Tilman N. Höhle*, 107–191. Berlin: Language Science Press. Repr. from *Satzglieder im Deutschen. Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*. Werner Abraham (ed.). Tübingen: Narr. 75–153.

Jeffreys, Harold. 1939. *Theory of probability*. Oxford: Clarendon Press.

Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. University of Edinburgh. (Doctoral dissertation).

Kiss, Tibor. 2012. Reflexivity and dependency. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax* (Linguistische Arbeiten 547), 155–185. Berlin: De Gruyter.

Kupietz, Marc, Cyril Belica, Holger Keibel & Andreas Witt. 2010. The German Reference Corpus DeReKo: A primordial sample for linguistic research. In *Proceedings of the 7th conference on International Language Resources and Evaluation (LREC 2010)*, 1848–1854.

Landau, Idan. 2010. *The locative syntax of experiencers*. Cambridge, MA: The MIT Press.

Lange, Kristian, Simone Kühn & Elisa Filevich. 2015. "Just another tool for online studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLOS ONE* 10(7). 1–14. DOI: 10.1371/journal.pone.0130834.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen* (Studien zur deutschen Grammatik 5). Tübingen: Gunter Narr.

Liddell, Torrin M. & John K. Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. DOI: 10.1016/j.jesp.2018.08.009.

Masloch, Simon, Johanna M. Poppek & Tibor Kiss. 2024. Not so peculiar after all: On the normal position of arguments of German experiencer-object verbs. *Glossa* 9. 1–37. DOI: 10.16995/glossa.10150.

Müller, Gereon. 1999. Optimality, markedness, and word order in German. *Linguistics* 37(5). 777–818. DOI: 10.1515/ling.37.5.777.

Nicenboim, Bruno, Daniel J. Schad & Shravan Vasishth. 2024. *Introduction to Bayesian data analysis for cognitive science*. https://vasishth.github.io/bayescogsci/book/.

Pesetsky, David Michael. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Pieper, Jutta, Alicia Katharina Börner & Tibor Kiss. 2025. Identifying non-cooperative participation in web-based elicitation of acceptability judgments: How to get rid of noise in your data. *Journal of Research Design and Statistics in Linguistics and Communication Science*. to appear.

Platzack, Christer. 2012. Backward binding and the C-T phase: A case of syntactic haplology. In Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *Functional heads: The cartography of syntactic structures*, 197–207. Oxford: Oxford University Press.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

R Core Team. 2023. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4). 657–720.

Reuland, Eric. 2011. *Anaphora and language design* (Linguistic Inquiry Monographs). Cambridge, MA: MIT press.

Reuland, Eric & Tanya Reinhart. 1995. Pronouns, anaphors and case. In Hubert Haider, Susan Olsen & Sten Vikner (eds.), *Studies in comparative Germanic syntax*, 241–268. Kluwer. DOI: 10.1007/978-94-015-8416-6_11.

Rozwadowska, Bożena, Arkadiusz Nowak & Anna Bondaruk. 2020. Psych verbs: Setting the scene. In Bożena Rozwadowska & Anna Bondaruk (eds.), *Beyond emotions in language: Psychological verbs at the interfaces*, 1–21. Amsterdam: John Benjamins. DOI: 10.1075/la.263.01roz.

Rychlý, Pavel. 2008. A lexicographer-friendly association score. In Petr Sojka & Aleš Horák (eds.), *Proceedings of recent advances in Slavonic natural language processing, RASLAN 2008*, 6–9. Brno: Masaryk University.

Scheepers, Christoph, Barbara Hemforth & Lars Konieczny. 2000. Linking syntactic functions with thematic roles: Psych-Verbs and the resolution of subject-object ambiguity. In Barbara Hemforth & Lars Konieczny (eds.), *German sentence processing*, 95–135. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-015-9618-3_4.

Temme, Anne & Elisabeth Verhoeven. 2016. Verb class, case, and order: A crosslinguistic experiment on non-nominative experiencers. *Linguistics* 54(4). 769–813. DOI: 10.1515/ling-2016-0018.

Temme, Anne & Elisabeth Verhoeven. 2017. Backward binding as a psych effect: A binding illusion? *Zeitschrift für Sprachwissenschaft* 36(2). 279–308. DOI: 10.1515/zfs-2017-0011.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

Webelhuth, Gert. 2022. C-command constraints in German: A corpus-based investigation. *Zeitschrift für Sprachwissenschaft* 41(2). 339–392. DOI: 10.1515/zfs-2022-2001.