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

Simon Masloch, Johanna M. Poppek, Tibor Kiss

Ruhr-University Bochum

May 27, 2024

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 if the antecedent precedes the reflexive in linear order and 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; Cheung & Larson, 2015; Landau, 2010; Pesetsky, 1995; Pollard & Sag, 1992). In Belletti and Rizzi (1988)'s 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.

(1)   Italian (Belletti & Rizzi, 1988, p. 312)

a. Questi pettegolezzi su di sé preoccupano Gianni più di ogni altra cosa.

'These gossips about himself worry Gianni more than anything else.'

b. * Questi pettegolezzi su di sé descrivono Gianni meglio di ogni biografia ufficiale.

'These gossips about himself describe Gianni better than any official biography.'

While 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) – the subject is taken to originate in a position *below* the object, so 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. Bouchard, 1995; Pollard & Sag, 1992).

In this paper, we will present evidence from an offline acceptability study that binding into the subject of EO verbs in German is possible only if the object precedes (and 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. Fischer, 2015; Kiss, 2012; Platzack, 2012; 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 and 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 *was* possible, a logophoric interpretation of the reflexive could not 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 in our experimental study 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* stands in need of 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 the 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. Together they will lead us to expectations about the acceptability of reflexive binding into the subjects of German EO verbs, which will be discussed there, too. 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 aspects of German syntax and some facts about "backward binding" with EO verbs that will be relevant for the discussion later on.

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), so 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 (Höhle, 2019/1982)) 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) (i.a. Frey, 1993; Haider, 2017). Other approaches favour base-generation of the different orders (i.a. Fanselow, 2001) or do 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 and 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. Ellsiepen & Bader, 2018; Fanselow, 1992; Haider &

Rosengren, 2003; Lenerz, 1977; Lötscher, 1981; Primus, 2004; Scheepers et al., 2000), recent experimental evidence points to a preference for object before subject with (almost all) dative-object EO verbs and a preference for subject before object with (almost all) accusative-object EO verbs if subjects are inanimate and all other factors potentially influencing linear order are controlled for (Masloch et al., under review; Temme & Verhoeven, 2016).

Masloch et al. (under review) assume base generation of sentence-level constituents and violable linear precedence constraints to capture 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 a much research on the factors influencing the linear order of elements in the German midfield (see i.a. Ellsiepen & Bader, 2018; Hoberg, 1997; Keller, 2000; Lenerz, 1977; Lötscher, 1981; Uszkoreit, 1987). As Masloch et al. (under review) 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 ≺ NON-CAUSER, which places causers before non-causers and has more weight than the constraint ANIMATE ≺ 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 following schematic structure emerges (the labelling is not important here):

(2)
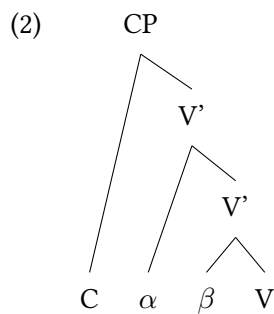
We will abstract away from certain phenomena, which are orthog-

onal to our analysis, such as the fronting of (reflexive) pronouns, the possibility to place constituents in the prefield, 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, but an approach assuming a fixed base order plus scrambling conceived as movement to the left (as e.g. in Haider, 2017) will basically produce the same predictions, as long as scrambling is not undone to reconstruct for anaphoric binding or binding constraints may 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; for Japanese Fujita, 1993). We already saw an example of this in (1a), repeated here.

(1a)   Italian (Belletti & Rizzi, 1988, p. 312)
       Questi pettegolezzi su di sé preoccupano Gianni più di ogni altra cosa.
       'These gossips about himself worry Gianni more than anything else.'

Such examples pose a problem for theories of binding that require an anaphor to be c-commanded by its antecedent. Many solutions have been proposed in the literature, which 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 verb class) (e.g. Belletti & Rizzi, 1988; Cheung & Larson, 2015; Pesetsky, 1995) and those that relate it to exemption form Principle A and logophoric or point-of-view-based binding, which is available more generally (e.g. Bouchard, 1995; Pollard & Sag, 1992). On Belletti and Rizzi' (1988) account of the syntax of psych verbs, the subject 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 assume the subject to originate in a position below the object are still widely assumed for dative-object EO verbs and sometimes for ((some readings

of) some) accusative-object EO verbs (see the overview in Rozwad-owska et al., 2020), using backward binding to argue for it has some-what fallen out of fashion (cf. e.g. Hirsch, 2018; Landau, 2010). This is so because various authors have shown that backward binding is even possible despite the impossibility to establish a c-command re-lationship between the antecedent and the putative anaphor at any given syntactic level (see e.g. Bouchard, 1995; Cançado & Franchi, 1999; Pollard & Sag, 1992). An illustrative example from Brazilian Portuguese given by Cançado and Franchi (1999) is (3). Logophoric binding can account for such cases, so one needs to assume it any-way.

(3)   Brazilian Portuguese (Cançado & Franchi, 1999, p. 140)
      Rumores sobre si   explicam a   insegurança mostrada
      rumors   about himself explain   the insecurity   shown
      por João.
      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 (4a), nor non-c-commanding (4b), nor split antecedents of embedded reflexives (4c). The examples in (4) are not only somewhat degraded but grossly unacceptable on the coindexations given.

(4)   (Kiss, 2012, 158, glosses adapted and translations added)
      a. * Ulrich$_1$ war sauer. Ein Bild   von sich$_1$ war
          Ulrich was upset a   picture of   REFL had
          beschädigt worden.
          damaged   been
          'Ulrich was upset. A picture of himself had been
          damaged.'
      b. * [Schumachers$_1$ Reklamevertrag]   verlangte eine
          Schumacher.GEN promotion.contract required   a
          Nacktaufnahme von sich$_1$.
          nude.photo   of   REFL
          'Schumacher's promotion contract required that nude
          photos of himself be taken.'
      c. * Ulrich$_1$ zeigte   Klaus$_2$ einige Bilder   von sich$_{1+2}$.
          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 ex-emption, it should thus *not* be possible in German. If it *is* possi-ble, 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. Fischer, 2015; Kiss, 2012; Platzack, 2012; 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 (5b) is acceptable to her.

(5) (Kiss, 2012, 161, glosses adapted and translations added; b. acceptable according to Fischer, 2015, p. 1390)

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

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

While Kiss (2012) judges (5b) 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 and Verhoeven (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 EO verbs, one for dative-object EO verbs), in which they compared the acceptability of backward binding into the subjects of EO verbs

and agentive verbs in two conditions: particular or 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 and 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 Temme and Verhoeven observe. 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, p. 387). In an article about argument linearisation, Temme and Verhoeven (2016) argue that experiencers are more likely aboutness-topics than patients. Thus, it may be the case that Temme and 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. For this reason, we consider it preferable to use 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 are 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 which either allows reconstruction or may be fulfilled at any or an early point. Both examples in (5) should be grammatical then. If arguments are base-generated in their surface positions or scrambling

does not reconstruct for binding, only (5b) should be grammatical. The latter view is argued by Masloch et al. (under review) to explain the linearisation preferences they observe with German EO verbs less naturally.

A potential problem for an experimental investigation into the possiblility of reflexive backward binding in German pointed out by Temme and Verhoeven (2017, p. 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 (6).

(6)   (based on an in-text example by Temme & Verhoeven, 2017, p. 286)
      Er betrachtete seine Möbel   / $^{??}$die Möbel   von sich.
      he beholded   his   furniture   the furniture of   REFL
      'He looked at his furniture.'

Such functional overshadowing may lead to reduced acceptability of the stimuli in an experimental setting, so stimuli where this may happen are to be avoided. This is possible since not all combinations of noun, preposition and reflexive are equally unacceptable: e.g. *Bilder von sich* 'pictures of REFL' as in (5) is not generally unacceptable as shown by sentences like *Warum hat Claude Cahun₁ die Bilder von sich₁ zurückgehalten?* 'Why has Claude Cahun withheld the pictures of herself?' (Kiss, 2012, p. 156).

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 pregistered 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 (7,8) (presented here without any acceptability judgment):

(7) Accusative object

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

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

    'It is obvious that the rumour about himself annoyed the professor.'

(8) Dative object

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

    b. Object before subject:
Es ist allgemein bekannt, dass dem   Opernsänger
it is commonly known   that the.DAT opera.singer.DAT
die   Meldung   über sich gefallen   hat.
the.NOM 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.[1] Verbs were chosen based on

---

[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 like', *imponieren* 'to impress', *missfallen* 'to displease', *nahegehen* 'to afflict', *widerstreben* 'to have an aversion'; 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'

their syntactic behavior in corpus data, essentially following and using the materials from Masloch et al. (under review), 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 (6)). In order to do so, nouns frequently having a preposition as its right neighbour were extracted from DeReKo (Kupietz et al., 2010) using KorAP (Diewald et al., 2016). From these, 327 nouns were manually chosen, then for each of them collocation scores with 81 prepositions (as direct right neighbours) and possessive pronouns (maximally three words to the left of the noun) were computed. We then chose noun-preposition combinations from the pairs with a high logDice (an association score defined by Rychlý (2008)) and 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.

In addition to the test items, questionnaires contained 64 filler items. Among them, there were six (unannounced) calibration items included to familiarise participants with the task, 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 (e.g., no two test items should follow each other).

3.3 Participants and Procedure

Since CASE was manipulated between items and ORDER within items, there were 16 test items in two linearisation conditions (*SO* and *OS*). We created two lists so that each participant rated only one ordering condition per item. Participants (monolingual native speakers of German, DACH-residents) 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 where the participants' individual reaction times were automatically measured. Taking the control- and attention checks specified in the pre-registration as well as possible topic awareness (checked with an open question at the end of the survey) into account, data from 48 participants has been included in the analysis.[2] After giving their informed consent to participate in the experiment, participants read written instructions asking them to rate how natural the 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. (under review)'s account of argument linearisation in the midfield for German EO verbs, which takes surface-order at face value: A constituent in the midfield is taken to c-command another if and only if it precedes it 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).[3]

(9)   Main hypothesis
      If the subject precedes the object in the target clause, the item is ungrammatical.

---

[2] The predefined exclusion criteria: 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 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 (2023) method, 7. the participant self-declared to reside 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 (2023) criteria for reaction times, which appeared to be too strict given the overall very fast reaction times. Overall, the 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. However, 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.

[3] 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.

	dative	accusative
OS	high	medium
SO	low/medium	low

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

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 unacceptability. An OS linearisation will violate CAUSER ≺ NON-CAUSER with (most) accusative-, but not with (most) dative-object EO verbs (see Section 2.1). An SO linearisation violates ANIMATE ≺ INANIMATE, which is outweighted 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 hypothetise that participants will know what the unmarked order would have been for a given sentence and that for at least some of them the answer provided may be influenced by this alternative, especially when they try to behave like cooperative discourse-participants (remember that we asked participants to rate how natural the sentences sound to them) when being confronted with an ungrammatical sentence. This may result in less than completely low ratings. 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.

We will use a Bayesian generalised linear mixed model (Bürkner, 2017) to analyse the data. Given a cumulative link generalised mixed model containing both factors and their interaction with *accusative* and *SO* coded −1, *dative* and *SO* 1 (sum-coding), our assumptions lead to the following expected effects:[4]

---

[4]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
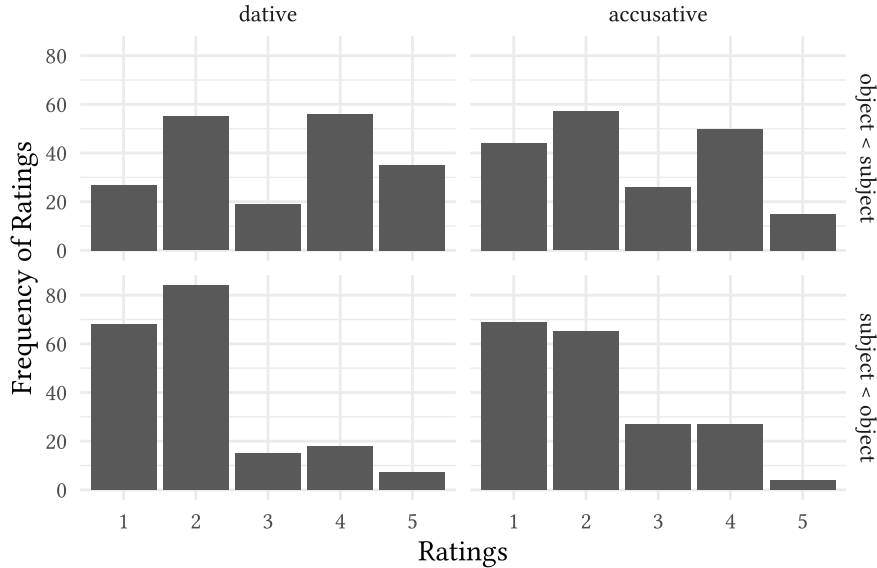
Figure 1: Empirical distribution of ratings. "5" stands for "completely natural", "1" for "completely unnatural". All choice options were presented to the participants with a natural language label.

(10) Expectations fixed effects sum-coded model:
- CASE: Mildly positive
- ORDER: Medium/strong positive
- CASE × ORDER: marginal or non-existent

3.5 Results

Figure 1 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 (9).

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). Both factors were sum-coded with *dative* and *OS* coded 1 and

---

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.

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.[5]

Bayesian cumulative generalised mixed models differ from vanilla linear regression models in several ways: In a mixed model, the strength of effects may vary per e.g. participant or item. In a generalised mixed model, a link function is applied to the linear predictor. Cumulative models are a type of model that can be used to analyse ordinal data (see Bürkner and 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). 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. The probability that a response option is chosen then depends on the linear predictor and the thresholds. We use a logit model, so the probability that a response option r of an ordered variable Y is chosen is modelled as $P(Y = r \mid \eta) = logistic(\tau_{r|r+1} - \eta) - logistic(\tau_{r-1|r} - \eta)$, where η is the linear predictor, $\tau_{r|r+1}$ is the threshold between r and the next higher response option, $\tau_{r-1|r}$ is the threshold between r and the next lower response option, and $logistic$ is the logistic distribution (for the first response option the subtrahend will be 0, for the last one, the minuend will be 1). Figure 2 shows the (point) estimates for the thresholds from the model reported below. While the estimates for the thresholds are model-dependent and thus uninteresting in themselves, it is obvious that the response options correspond to portions of the latent variable of unequal size. A metric model would treat the distances as being equal.

In Bayesian models, parameters are random variables, so one can talk about the credibility of different values (an introductions 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. Inferences can then be based on the posterior, 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-

---

[5]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.

completely
unnatural   rather unnatural   no tendency   rather natural   completely
natural

-2 -1 0 1 2 3 4

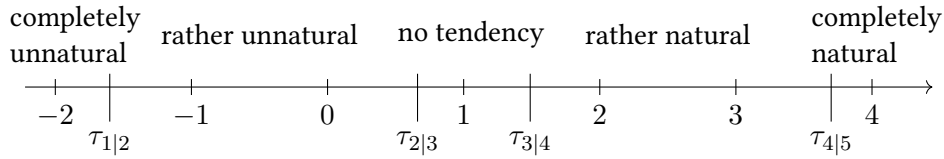$\tau_{1|2}$ $\tau_{2|3}$ $\tau_{3|4}$ $\tau_{4|5}$

Figure 2: Point estimates for thresholds and proportions of the latent variable corresponding to the response options.

neutral regularising priors.[6] 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 (see Nicenboim et al. 2024 chapter 15 for an introduction). We will compare the model we report here against models 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.[7]

There is an effect of ORDER in the expected direction ($\hat{\beta} = 0.79$, 95 % credible interval (CrI) $= [0.47, 1.12]$)[8], which is, however, not quite as strong as expected as it roughly corresponds to the difference between two levels of the ordered response variable.[9] Nevertheless, the Bayes factor shows that there is extreme evidence for this effect (Bayes factor computed as Savage-Dickey density ratio between the model and a model where β is set to 0 (BF_{10}) $= 126.7$). The point estimate for the effect of CASE ($\hat{\beta} = 0.17$, CrI $= [-0.24, 0.57]$, $\text{BF}_{10} = 0.074$) 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 in unmarked for dative-object EO verbs, but not for the accusative-object ones. While descriptively sentences are rated higher in *dative*

---

[6]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.

[7]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. For our theoretical purposes, the results are similar.

[8]We use hats (ô) 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).

[9]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.

OS than in *accusative OS*, it is the other way round for the *SO* sentences. Judgments of naturalness may correspond to normal order in a more straightforward, 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 ($\text{BF}_{10} = 0.138$). Again, if it is there, it is relatively small.

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 also not small ($\hat{SD} = 0.5$, CrI $= [0.26, 0.75]$), while the ones for CASE ($\hat{SD} = 0.16$, CrI $= [0.01, 0.39]$) and the interaction ($\hat{SD} = 0.22$, CrI $= [0.01, 0.48]$) are. There is variation between items ($\hat{SD} = 0.74$, CrI $= [0.45, 0.1.18]$), 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. This may have to do with the rather high complexity of the items and the relative scarcity of PPs containing a reflexive embedded in an NP in combination with the fact that we asked participants to rate how natural the sentences sound. 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

The results support our main predictions: Reflexive binding into the subject of German EO verbs is licit only if it is *not* backward.[10] 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 *not* if EO verbs are taken to be unaccusative (which should translate into an OS base order) and scrambling is taken 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. If we take the unaccusativity hypothesis to imply that the base order of the respective verbs is object before subject, our data are compatible with it if scrambling in German does not reconstruct for binding, but then they are equally compatible with an SO base order. Both a base-generation approach as descibed 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 however. 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. Rather, our model suggests that there is a slight tendency for participants to rate an item in which the order of subject and object is unmarked higher than one in which it is not irrespective of overall grammaticality (interaction effect), even though we cannot be sure about its existence. If it is there, it is small and we may suspect it to be reducible to performance mistakes.

Two other remarkable aspects are the unexpectedly low level of acceptability and the strong individual variation. 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. Thus, some test items may appear artificial to some participants, resulting in differing ratings of their "naturalness". It is also well conceivable that participants differ in how natural the noun-preposition-reflexive combinations sound to them, which may have an influence on the naturalness of the embedded reflexive be-

---

[10] One may object that both orderings could be ungrammatical and the enhanced acceptability of the OS order is due to priming effects. An account along such lines is not tenable since the examples in (4), where the possible antecedent precedes the reflexive, too, are clearly unacceptable.

cause the overshadowing process discussed in Section 2.2 presumably involves competition between a PP containing the reflexive and the possessive. Participants may weigh the factors involved in the competition differently, resulting in overall different judgments of naturalness. 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 bad. 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 lowish 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 more narrow. However, given the availability of a plausible alternative, we do not want to pursue this path.

4.2 Theories of binding

So far, we assumed that the German reflexive *sich* has to be c-commanded by its antecedent, which is in line with classical binding theory's Principle A (Chomsky, 1981). 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 and Reuland' (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 and Reinhart (1995). Pollard and 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 and Reuland (1993) and condition A does not apply, and the anaphor does not have a less oblique co-argument so that Pollard and Sag's (1992) Principle A does not apply. Thus, the reflexive should be licensed without being bound.[11] Even if there was 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

---

[11] Pollard and 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.
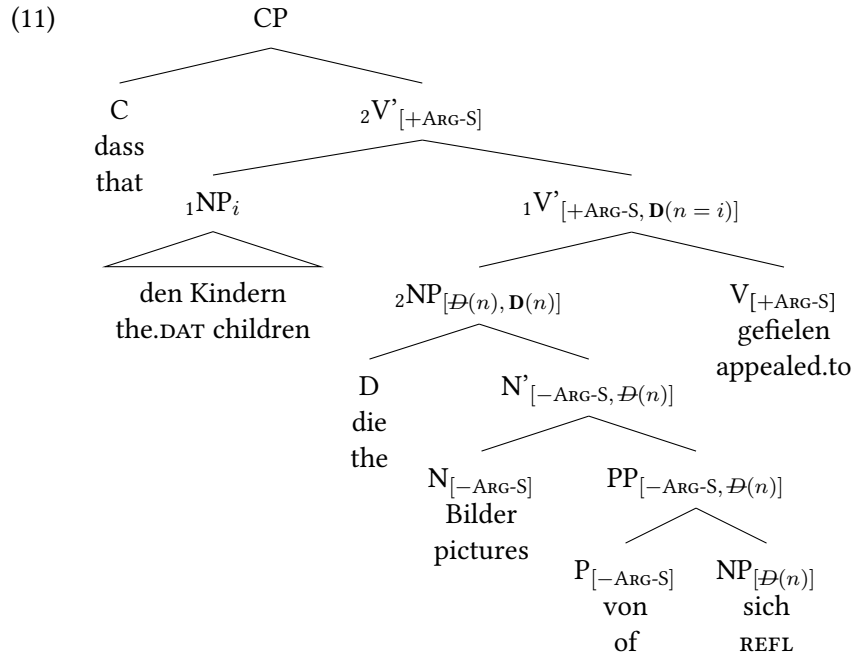
could not capture our data, since it is the positioning of the whole syntactic predicate (the NP/PP) that makes the difference, so some structural condition must be at play.[12]

That being said, the 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 $\mathbf{D}(n)$, where n is an index, if it is an argument of a head with an articulate argument structure (a verb or event noun; bearing the feature $+$Arg-S). $\mathbf{D}(n)$ is projected upwards until n is identified with the index of a sister of a phrase bearing the feature. In German, a feature $\mathbf{Đ}(n)$ representing an inactive dependency is introduced if the head does not have an articulate argument structure ($-$Arg-S), and projected in the same way until it is activated (= leads to the phrase having the feature $\mathbf{D}(n)$) when meeting a head with an articulate argument structure.[13] 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.[14] (11) illustrates how this works for (5b). Since P and N in (11) do not have an articulate argument structure, the $\mathbf{Đ}(n)$ introduced with *sich* becomes active only once $_2$NP – we use natural numbers to distinguish nodes in the tree and small letters for indices – becomes a daughter of the verbal projection, i.e. $_2$NP bears $\mathbf{D}(n)$. $\mathbf{D}(n)$ 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 (5a) and our SO items, leading to ungrammaticality.

---

[12]According to Reuland and 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, pp. 249 sqq.; Reuland, 2011, pp. 275 sqq.), but in our judgment stressed *sich* is perfectly fine in examples like (5b). 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 because Reuland (2011, pp. 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.

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

[14]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'.

(11)

$$\text{CP}$$
$$\text{C / }_2\text{V'}_{[+\textsc{Arg-S}]}$$

C
dass
that

$_2\text{V'}_{[+\textsc{Arg-S}]}$

$_1\text{NP}_i$

den Kindern
the.DAT children

$_1\text{V'}_{[+\textsc{Arg-S}, \mathbf{D}(n=i)]}$

$_2\text{NP}_{[\cancel{D}(n), \mathbf{D}(n)]}$

$\text{V}_{[+\textsc{Arg-S}]}$
gefielen
appealed.to

D
die
the

$\text{N'}_{[-\textsc{Arg-S}, \cancel{D}(n)]}$

$\text{N}_{[-\textsc{Arg-S}]}$
Bilder
pictures

$\text{PP}_{[-\textsc{Arg-S}, \cancel{D}(n)]}$

$\text{P}_{[-\textsc{Arg-S}]}$
von
of

$\text{NP}_{[\cancel{D}(n)]}$
sich
REFL

5 Conclusion

Our study shows that reflexive binding into the subject of experiencer-object verbs in 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 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

Abbreviations

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

References

Belletti, A., & Rizzi, L. (1988). Psych-verbs and θ-theory. *Natural Language & Linguistic Theory*, *6*(3), 291–352. https://doi.org/10.1007/BF00133902

Bouchard, D. (1995). *The semantics of syntax: A minimalist approach to grammar.* The University of Chicago Press.

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, *2*(1), 77–101. https://doi.org/10.1177/2515245918823199

Cançado, M., & Franchi, C. (1999). Exceptional binding with psych-verbs? *Linguistic Inquiery*, *30*, 133–143.

Cheung, C. C.-H., & Larson, R. K. (2015). Psych verbs in English and Mandarin. *Natural Language & Linguistic Theory*, *33*(1), 127–189. https://doi.org/10.1007/s11049-014-9259-3

Chomsky, N. (1981). *Lectures on Government and Binding.* Foris Publications.

Diewald, N., Hanl, M., Margaretha, E., Bingel, J., Kupietz, M., Bański, P., & Witt, A. (2016). KorAP architecture: Diving in the deep sea of corpus data. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, 3586–3591.

Ellsiepen, E., & Bader, M. (2018). Constraints on argument linearization in German. *Glossa*, *3*(1), 1–36. https://doi.org/10.5334/gjgl.258

Fanselow, G. (1992). „Ergative" Verben und die Struktur des deutschen Mittelfelds. In L. Hoffmann (Ed.), *Deutsche Syntax: Ansichten und Aussichten* (pp. 276–303). de Gruyter.

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

Fischer, S. (2015). Theories of binding. In T. Kiss & A. Alexiadou (Eds.), *Syntax – theory and analysis: An international handbook* (pp. 1357–1400). De Gruyter Mouton. https://doi.org/10.1515/9783110363708-016

Frey, W. (1993). *Syntaktische Bedingungen für die semantische Interpretation.* Akademie Verlag.

Frey, W. (2006). Contrast and movement to the German prefield. In V. Molnár & S. Winkler (Eds.), *The Architecture of Focus* (pp. 235–264). Mouton de Gruyter.

Fujita, K. (1993). Object movement and binding at LF. *Linguistic Inquiry, 24*(2), 381–388.

Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a Maximum Entropy model. In J. Spenader, A. Eriksson, & Ö. Dahl (Eds.), *Proceedings of the Stockholm workshop on 'variation within Optimality Theory'* (pp. 113–122).

Grewendorf, G., & Sabel, J. (1999). Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language and Linguistic Theory, 17*, 1–65. https://doi.org/10.1023/A:1006068326583

Haider, H. (2010). *The Syntax of German.* Cambridge University Press. https://doi.org/10.1017/CBO9780511845314

Haider, H. (2017). Mittelfeld phenomena: Scrambling in Germanic. *The Wiley Blackwell companion to syntax* (2nd ed., pp. 2573–2645). Wiley. https://doi.org/10.1002/9781118358733.wbsyncom048

Haider, H., & Rosengren, I. (2003). Scrambling: Nontriggered chain formation in OV languages. *Journal of Germanic Linguistics, 15*(3), 203–267. https://doi.org/10.1017/S1470542703000291

Hirsch, N. (2018). *German psych verbs – insights from a decompositional perspective* (PhD Thesis). Humboldt-Universität zu Berlin.

Hoberg, U. (1997). Die Linearstruktur des Satzes. In G. Zifonun, L. Hoffmann, & B. Strecker (Eds.), *Grammatik der deutschen Sprache* (pp. 1495–1680). De Gruyter.

Höhle, T. N. (2019/1982). Explikationen für „normale Betonung" und „normale Wortstellung" [Original publication 1982 in Werner Abraham (ed.). Satzglieder im Deutschen. Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung, 75–153. Tübingen: Narr]. In S. Müller, M. Reis, & F. Richter (Eds.), *Beiträge zur deutschen Grammatik: Gesammelte Schriften von Tilman N. Höhle* (2nd ed., pp. 107–191). Language Science Press.

Jeffreys, H. (1939). *Theory of probability.* Clarendon Press.

Keller, F. (2000). *Gradience in Grammar : Experimental and Computational Aspects of Degrees of Grammaticality* (Doctoral dissertation). University of Edinburgh. Edinburgh.

Kiss, T. (2012). Reflexivity and dependency. In A. Alexiadou, T. Kiss, & G. Müller (Eds.), *Local modelling of non-local dependencies in syntax* (pp. 155–185). De Gruyter.

Kupietz, M., Belica, C., Keibel, H., & Witt, A. (2010). The German Reference Corpus DeReKo: A primordial sample for linguistic research. *Proceedings of the 7th conference on International Language Resources and Evaluation (LREC 2010)*, 1848–1854. http://www.lrec-conf.org/proceedings/lrec2010/pdf/414_Paper.pdf

Landau, I. (2010). *The Locative Syntax of Experiencers.* MIT Press.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen.* Gunter Narr.

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*, 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

Lötscher, A. (1981). Abfolgeregeln für Ergänzungen im Mittelfeld. *Deutsche Sprache*, *9*, 44–60.

Masloch, S., Poppek, J. M., & Kiss, T. (under review). *Not so peculiar after all: On the normal position of arguments of German experiencer-object verbs.* https://ling.auf.net/lingbuzz/007118

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

Nicenboim, B., Schad, D. J., & Vasishth, S. (2024). *Introduction to Bayesian data analysis for cognitive science.* https://vasishth.github.io/bayescogsci/book/

Pesetsky, D. M. (1995). *Zero syntax: Experiencers and cascades.* MIT Press.

Pieper, J., Börner, A. K., & Kiss, T. (2023). *Identifying non-cooperative participation in web-based elicitation of acceptability judgments: How to get rid of noise in your data.* https://ling.auf.net/lingbuzz/006514

Platzack, C. (2012). Backward binding and the C-T phase: A case of syntactic haplology. In L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, & C. Poletto (Eds.), *Functional heads: The cartography of syntactic structures* (pp. 197–207). Oxford University Press.

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

Primus, B. (2004). Protorollen und Verbtyp: Kasusvariation bei psychischen Verben. In R. Kailuweit & M. Hummel (Eds.), *Semantische Rollen* (pp. 377–401). Gunter Narr.

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

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

Reuland, E. (2011). *Anaphora and language design.* MIT press.

Reuland, E., & Reinhart, T. (1995). Pronouns, anaphors and case. In H. Haider, S. Olsen, & S. Vikner (Eds.), *Studies in comparative Germanic syntax* (pp. 241–268). Kluwer.

Rozwadowska, B., Nowak, A., & Bondaruk, A. (2020). Psych verbs: Setting the scene. In B. Rozwadowska & A. Bondaruk (Eds.), *Beyond Emotions in Language: Psychological verbs at the interfaces* (pp. 1–21). John Benjamins. https://doi.org/10.1075/la.263.01roz

Rychlý, P. (2008). A lexicographer-friendly association score. In P. Sojka & A. Horák (Eds.), *Proceedings of recent advances in slavonic natural language processing, RASLAN 2008* (pp. 6–9). Masaryk University.

Scheepers, C., Hemforth, B., & Konieczny, L. (2000). Linking syntactic functions with thematic roles: Psych-Verbs and the resolution of subject-object ambiguity. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 95–135). Springer Netherlands. https://doi.org/10.1007/978-94-015-9618-3_4

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

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

Uszkoreit, H. (1987). *Word Order and Constituent Structure in German.* CSLI Publications.

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