

Inverse Case attraction: experimental evidence for a syntactically guided process

Anna Czypionka^{1,2} Laura Dörre¹
Josef Bayer¹

Abstract In progressive Case attraction, the Case of a head nominal overwrites the Case of a following coindexed relative pronoun. The reverse process is called ‘inverse’ Case attraction. There, the morphologically overt Case of a relative pronoun overwrites the Case of a preceding head nominal. Inverse Case attraction has been attested in languages like Ancient Greek, Latin, and in the history of different Germanic languages. For modern standard German, its existence has in general been denied. We first discuss current analyses which have nevertheless identified inverse Case attraction in modern German on the basis of historical data and experimental judgement studies. We then present four behavioral experiments on the processing of German sentences. Effects of inverse Case attraction in the comprehension of German are revealed in self-paced reading times. They are fundamentally different in structures allowing attraction of dative Case than in structures allowing attraction of accusative Case, with much stronger effects for dative than for accusative Case. The results are interpreted in a theory of Case that draws a syntactic difference between structural and inherent (‘lexical’) Case rather than along the lines of the familiar Case hierarchy.

Keywords Case · Inverse Case attraction · Dative · German · Self-paced reading

Anna Czypionka
anna.czypionka@uni-konstanz.de

Laura Dörre
laura.doerre@uni-konstanz.de

Josef Bayer
josef.bayer@uni-konstanz.de

¹ Department of Linguistics, University of Konstanz, Fach D 185, 78457 Constance, Germany

² Center for Experimental Research on Natural Language, University of Wrocław, Wrocław, Poland

1 Introduction

Case attraction is a process in which the Case of a nominal constituent is copied onto another nominal constituent whose actually expected Case is overwritten. The Case of the donor is always more marked than the Case of the receiver. This process seems to be confined to the interaction between head nominal and the following relative pronoun of a relative clause, requiring adjacency, coindexation and morphologically overt unambiguous Case marking to occur.

The notion *Case attraction* is classically understood as *progressive Case attraction*, i.e., the transfer from the head nominal to the relative pronoun. There is a second notion, *inverse Case attraction*, which refers to a reverse process in which the relative pronoun donates its Case to the head nominal. Within the Germanic languages, Harbert (2007: 423) describes inverse attraction as a rare process which is occasionally found in Old English and Old (OHG), Middle (MHG) and Early New High German (ENHG). Case attraction appears to play no role in the production of standard German, a language in which enough Case morphology has survived to make the effect visible.

The current article nevertheless sets out to shed light on inverse Case attraction in modern German. We will present an analysis of the construction allowing inverse Case attraction to occur, and will provide experimental evidence for this type of Case attraction in comprehension. It will show that these effects are not confined to the constructions that have been held responsible for its emergence since the nineteenth century. The traditional literature treats Case attraction as a purely morpho-syntactic phenomenon. Language processing does not come into consideration. This is surprising insofar as the classical examples of attraction are optional and in fact are the exception rather than the rule. Straight syntactic rules operate mandatorily. We will see below that there are good reasons to reconsider at least core cases of attraction in the light of linguistic performance rather than as a purely grammatical phenomenon. At the same time, we will show that performance is strictly under the control of linguistic structure.

In Sect. 2, we first give an overview of the theoretical questions on Case attraction, and elaborate our own assumptions about the phenomenon. We then continue with an overview of earlier experimental work on inverse Case attraction in comprehension, and contrast it with experimental findings generally attributed to successful processing of lexical Case marking. This section ends with the presentation of our own psycholinguistic research questions. In Sect. 3, we present the stimulus materials used in our studies. In Sects. 4 and 5, we present the results of four experiments. In Sect. 6, we discuss our experimental findings together and offer a theoretical interpretation.

2 Background

2.1 Case attraction in grammar

Case attraction is a well-known although not too well-studied phenomenon that can be found in many languages. A remarkable fact about Case attraction is that it has been observed quasi exclusively in the relation between an NP/DP and the relative operator of an immediately following relative clause. An NP/DP which bears a certain Case which we call CASE-1 passes CASE-1 on to the relative pronoun of a following relative clause. The established restriction is that CASE-1 can be attracted to the relative pronoun's CASE-2 only if CASE-1 is higher on the Case hierarchy than CASE-2, the assumed hierarchy being Nom < Acc < Dat < Gen. Acc or Dat or Gen can, for instance, overwrite Nom; Dat or Gen can also overwrite Acc; but Nom could never overwrite Acc or Dat or Gen; Acc could never overwrite Dat or Gen, etc.

Case attraction of this sort has been observed in various languages, according to Grimm (1866) in Ancient Greek, Latin, Gothic, Old Saxon, Old, Middle and New High German. In German, it is often assumed to have existed in older varieties but not in the modern standard language (cf. Pittner 1995, 1996). The following examples are taken from Pittner (1996) and the references show in her work.

(1) GEN instead of NOM

sie gedâht' ouch maniger leide, der ir dâ
she remembered also various sufferings.GEN which.GEN her at
 héimé geschach
home occurred
 'She also remembered various sufferings which she had to endure at home.'
 (MHG: Nib. 1391,14; Behaghel 1928: 756; Lenerz 1984: 116)

(2) DAT instead of NOM

sendida mih [...] zi **dheodom, dhem** euuuih biraubodon
sent me to the.one.DAT who.DAT you robbed
 'sent me to the one who robbed you'
 (OHG: Isid. 218f.; Helgander 1971: 174; Lenerz 1984: 116)

(3) GEN instead of ACC

alles des ich i.e. gesach
all.GEN which.GEN I ever saw
 'all I have ever seen'
 (MHG: Nib. 1698,1; Behaghel 1928: 756)

(4) ACC instead of NOM

unde ne wolden níet besên **den** **mort** **den**
and NEG wanted not take.notice the murder.ACC which.ACC
 dô was geschên
there has happened
 ‘and did not want to take notice of the murder that has happened there’
 (MHG: Alex. 3228; Grimm 1866: 319)

There is a reverse process in which it is the Case of the relative pronoun that is attracted to the ‘head’ NP/DP. This process is known as ‘inverse Case attraction’. Numerous examples from older stages of German can be found in Grimm (1866). (5a, b) show the dominant pattern in which nominative Case is replaced by the accusative. The pattern in (6) in which accusative is replaced by dative is rare. Nevertheless, the generalization appears to be correct that in all examples of inverse attraction CASE-1 can be attracted to CASE-2 only if CASE-1 is higher on the Case hierarchy than CASE-2.

(5) ACC instead of NOM

- a. **Den** **schilt** **den** er vür bôt **der** wart schiere
the shield.ACC which.ACC he held this.NOM was quickly
zeslagen.
hit.to.pieces
 ‘The shield that he held was quickly hit into pieces!’
 (MHG: Iwein 6722f.; Lernerz 1984: 116)
- b. **Den** **liebsten** **bulen** **den** ich hab **der** leit beim
the dearest lover.ACC who.ACC I have he.NOM lies at.the
wirt im keller.
inn.keeper in.the cellar
 ‘The dearest lover I have, he lies in the inn-keeper’s cellar.’
 (ENHG: Uhl. Volkslied 585, cited in Grimm 1866: 330)

(6) DAT instead of ACC

dem **gote** **dem** ich da dienen sol, **den** enhelfent
The god.DAT who.DAT I there serve should him.ACC NEG.help
si mir niht
they me not
 ‘the god whom I should serve, they don’t help me to [...] him’
 (MHG: MF 181, 25f.)

The fact that the Case on the ‘head’ NP/DP derives from the relative pronoun is proven by the severe ungrammaticality that results when the relative clause is deleted. Consider the following version of (5b).

(7) *den liebsten bulen der leit beim wirt im keller

(7) would amount to an example of *accusativus pendens*, a construction type that seems to be thoroughly unattested.

Despite the markedness hierarchy that governs Case attraction in both forms of attraction, some researchers doubt that inverse, i.e., regressive attraction should be seen on a par with regular, i.e., progressive attraction. Pittner (1995, 1996) observes that the relative clause is resumed by a pronoun which appears in the Case that is required by the matrix clause. From this, she concludes that inverse attraction should not be compared with regular attraction, for which such a restriction does not exist. Georgi and Salzmann (2014, 2017) come to a similar conclusion. In their minimalist probe-goal account of regular Case attraction, it is important that Case checking proceeds in top-down fashion. In the Swiss German data, that they have primarily in mind, the Case of the noun of the matrix-DP checks the unvalued Case on a zero relative operator in SpecCP. The copy of this operator in the VP can be valued by a Case-assigner if the feature set of the Case-assigner is a subset of the features of the Case-assignee. This process is called *Matching*. Matching does not require featural identity but at least a relation by which the Case-features required by the Case assigner, i.e. normally the verb, are a proper subset of the Case features on the zero operator. The familiar Case hierarchy is conceived as ascending in terms of the number of features.¹ A *prima vista* advantage of Georgi and Salzmann's model is that it brings syntactic derivations more directly in line with the online nature of human language processing.² Given that checking/valuation proceeds from left to right in a strictly downward branching structure, inverse attraction has no place in their system. Their conclusion is that progressive and inverse attraction are two completely different processes. They propose that inverse attraction crucially involves the correlative/left-dislocation that Pittner had already drawn our attention to. Recall that in (5) and (6) the left-dislocated part in which inverse attraction is observed does not interact with the main clause that follows.

In the following, we will give a different account of inverse Case attraction. We will concentrate on three issues:

- (i) Is Case attraction best explained with a Case hierarchy, or with a categorical distinction between lexical and structural cases?
- (ii) Which role does overt morphological marking play in Case attraction?
- (iii) Is Case attraction really confined to left-dislocation structures with resumptive pronouns, as proposed by Pittner (1995, 1996) and Georgi and Salzmann (2014, 2017)?

As for the first issue, it seems to be established that CASE-1 cannot be attracted to CASE-2 if CASE-1 is lower on the Case hierarchy than CASE-2. Given that it can

¹ Following Béjar and Āezáč (2009), Nom is $\{\alpha\}$, Acc is $\{\alpha\beta\}$, Dat is $\{\alpha\beta\gamma\}$, and Gen is $\{\alpha\beta\gamma\delta\}$. The system predicts equal gaps between the Cases. See also Caha (2009) for a nano-syntactic account of Case.

² Forerunners of such attempts are found in Phillips (2003) and Bianchi and Chesi (2014) among others.

never be the other way around, the question is still whether all the marked Cases are equally justified to overwrite any less marked Case, or whether there are differences by which some marked Cases enjoy a special privilege of being attracted. The Accessibility Hierarchy suggested by Keenan and Comrie (1977) as well as modern work on Case such as Caha (2009) suggest an obliqueness continuum. The generative tradition that started with the Government and Binding framework of Chomsky (1981) does not automatically deny such a continuum but predicts a split between the structural Cases nominative and accusative as a natural class and the verb-governed lexical Cases dative and genitive as another.

There is plenty of evidence that nominative and accusative as the two structural Cases form a natural class which patterns together whereas dative and genitive display deeply different behavior.³ Bayer et al. (2001) is a rich source of data and tests that demonstrate the deep split between structural and lexical Case, a split that seems to be superimposed on the Case hierarchy. The Case hierarchy is more morphology-oriented and suggests an equidistance in markedness between the Cases while the syntax of Case detects a split in this hierarchy by which Nominative and Accusative form a natural class that separates the two from the rest of the Cases. The two approaches are not incompatible with each other. Rather they focus on different aspects of the Case system.

The distinctness of the dative can be seen in widely discussed data from Swiss German relative clauses. Following van Riemsdijk (1989), Georgi and Salzmann (2014, 2017) report the standard pattern of Swiss relatives, namely that the structural Cases require a gap (which is bound by a zero operator in the specifier of the complementizer *wo*) whereas the dative (of the indirect object)—the genitive plays no role in Swiss German—must be spelled out as a resumptive pronoun as seen in (8c).

³ We hasten to say that we talk only about ad-verbal, not ad-nominal genitive. As one knows, ad-nominal genitive corresponds to the structural object Case but not to dative Case.

- (i) a. Ich sehe den Berg. b. Ich helfe dem Kranken.
 I see the.ACC mountain I help the.DAT patient
- (ii) a. die Sicht des Berges. b. *die Hilfe des Kranken.
 the sight the.GEN mountain the help the.GEN patient

(iib) does have an interpretation, but it could only mean that the patient is the agent, never that it is the beneficiary.

- (8) a. Ich suech de Bueb, wo (*er) immer z' spaat chunt_{NOM}.
I search the boy C he always too late comes
 'I'm looking for the boy who is always late.' SUBJECT
- b. Ich hilf em Bueb, wo-n -I (*en) geschter g'see_{ACC} han.
I help the boy C- N-I him yesterday seen have
 'I help the boy who I saw yesterday.' DIRECT OBJ.
- c. Das isch de Bueb, wo mer *(em) es Buech g'gee_{DAT} hand.
this is the boy C we him the book given have
 'This is the boy who we gave a book to.' INDIRECT OBJ.
- d. Das isch de Bueb, wo-n-i (*en) geschter gsee_{ACC} han.
This is the boy C-N-I him yesterday seen have
 'This is the boy who I saw yesterday.' DIRECT OBJ.

Obviously, NOM (as in 8a) and ACC (as in 8b) can be retrieved from the syntactic environment. There is no need to replace the trace with a resumptive pronoun. Therefore, economy prevents the choice of a resumptive. DAT (as in 8c), on the other hand, cannot be retrieved by formal means. It has to be inserted with its overt Case morphology. Insertion of the resumptive has the flavor of a repair strategy (cf. Bayer and Salzmann 2013).

In Georgi and Salzmann's account, *Bueb* in (8a) has an undischarged feature [**Acc**] and probes the zero relative pronoun to the left of the complementizer *wo* which has a potentially corresponding feature [*uAcc*]. After this, probe and goal have the same Case. The zero operator is then lowered to its argument position where it would actually be assigned Nom by the T-head related to *chunt*. A Case conflict emerges but can be resolved by Matching. The Case probe [**Nom**] on T can be discharged because its feature set is a subset of the set of features that has been established on the zero operator: $\{\alpha\} \subset \{\alpha\beta\}$. The same obtains in (8b) where the zero pronoun inherits dative Case while the verb *g'see* requires accusative. Matching is successful because of $\{\alpha\beta\} \subset \{\alpha\beta\gamma\}$. In (8c), however, matching cannot apply because the feature set probed by the dative-requiring verb *gee* is not a subset of the feature set on the zero pronoun: $\{\alpha\beta\gamma\} \not\subset \{\alpha\}$. In this case, a repair strategy applies by which a Case-bearing pronoun is inserted which was not in the numeration, and which gets bound by the zero operator. This is possible because zero operator and inserted pronoun agree in phi-features while they disagree in Case features.⁴

So far so good. A problem emerges though with example (8d). (8d), (25) of Georgi and Salzmann (2017), is surprisingly grammatical with a gap and ungrammatical with a resumptive. This is unexpected because the matrix Case which is passed to the zero operator is nominative, and the accusative assigned by *gsee* is not a subset of it: $\{\alpha\beta\} \not\subset \{\alpha\}$. The solution which Georgi and Salzmann offer

⁴ The psycholinguistic interpretability of the top-down derivation becomes less convincing at the point where matching comes into the picture. Case attraction gives Case to the zero relative pronoun right next to the Case probe, but then the relative pronoun can be linearly very far from its trace position. This means that the process of matching must be delayed. In processing terms, it must be kept in limbo until the end of the sentence is reached.

comes across like a stipulation. Case morphology plays close to no role in their system. At this point, however, they resort to the fact that nominative and accusative frequently coincide morphologically.⁵ Their suggestion is that nominative and accusative “represent the same type of Case, viz. unmarked Case”, and that they are therefore equally distinct from the dative. From the viewpoint of the markedness scale of the Case hierarchy on which their account rests, this relaxation is a severe deviation. It brings the scalar approach close to the distinction between structural and lexical Case and the fact that very often only the latter relies on overt Case morphology. Case distinctness by morphology is widely attested. In their work on Case-matching in free relatives, Groos and van Riemsdijk (1981) show that Case-ambiguous, syncretic, forms can cover different grammatical functions.

- (9) a. Was dort steht_{NOM} ist_{NOM} geschmacklos.
what there stands is tasteless
 ‘What is standing over there is tasteless.’
- b. Was du gekauft_{ACC} hast ist_{NOM} geschmacklos.
what you bought have is tasteless
 ‘What you bought is tasteless.’
- c. *Wen du eingeladen_{ACC} hast ist_{NOM} ein Idiot.
who.ACC you invited have is an idiot
 ‘Who you invited is an idiot.’
- d. *Wem du begegnet_{DAT} bist ist_{NOM} ein Idiot.
who.DAT you encountered are is an idiot
 ‘Who you ran into is an idiot.’
- e. Wem du begegnet_{DAT} bist würde kaum einer vertrauen_{DAT}.
who.DAT you encountered are would hardly anyone trust
 ‘Who you encountered, hardly anyone would trust.’

Both (9a) and (9b) are fine because the Case-syncretic neuter form *was* is compatible with nominative Case as in (9a) and with accusative Case as in (9b). (9c, d) are ill-formed, however, because the unambiguous (human) forms *wen* and *wem* are only compatible with accusative and dative Case respectively but not with the nominative Case that is required by the finite verb *ist* of the main clause. Case matching with two datives is observed in (9e).⁶ The *wh*-CP is the phrase that

⁵ While this is generally true in feminine, neuter and plural, it is not true in the masculine singular pronoun, not even in Swiss German where we find Nom *er* and Acc *en*. It may be important to know that examples that correspond to (8d) in Bavarian are ungrammatical with a zero relativizer and only grammatical with the overt relative pronoun *den* in the specifier of *wo* (cf. Bayer 1984).

(i) Des is der Bua, *(den) wo i gestern gseng_{ACC} hob.
this is the boy him C I yesterday seen have
 ‘This is the boy who I saw yesterday.’

⁶ This is exactly parallel to the examples that Georgi and Salzmann (2017) present in their (4).

satisfies the Case requirement of the verb *vertrauen*. The simplest theory would be one by which the *wh*-CP [*wem du begegnet bist*] is a dative-marked XP which has nominal features and is therefore allowed to bear dative Case. This makes it available to serve as the dative argument in the matrix clause's sub-structure ... *vertrauen*.

This shows that morphology is centrally involved in the process of Case assignment. Case-underspecified (syncretic) forms are ready to represent whatever Case is compatible with their form. This does not mean that unambiguous Case forms can never be overwritten as in fact is the case in (1) through (4). It means, however, that this is not what we expect in general. Hence the remarkability of Case attraction.

The third question is whether regressive Case attraction really is confined to left-dislocation structures with resumptive pronouns. Experimental evidence by Bader et al. (2000, 2001) and Bader and Bayer (2006) strongly suggests that inverse Case attraction in comprehension is not exclusively dependent on the correlative constructions as suggested by Pittner and Georgi and Salzmann. Rather, Case attraction in comprehension appears naturally and quite robustly also in sentences in which a Case-underspecified nominal is followed by an overtly Case-marked coindexed pronominal bearing dative Case. An example of stimulus material eliciting Case attraction is given in (10) (taken from Bader and Bayer 2006: 122).

- (10) a. #Ich weiß, dass Maria, der ich gerade
I know that Maria.NOM who.DAT I just
 begegnet bin, ein Päckchen geschickt hat.
encountered am a parcel sent has
 'I know that Maria, who I just ran into, has sent a parcel.'
- b. Ich weiß, dass Maria, der ich gerade
I know that Maria who.DAT I just
 begegnet bin, ein Päckchen geschickt wurde.
encountered am a parcel sent was
 'I know that Maria, who I just ran into, was sent a parcel.'

The embedded CP in (10a) must be Subject-Object-Verb (SOV) with *Maria* being the nominative subject. The embedded CP in (10b), however, must be Object-Subject-Verb (OSV) with *Maria* being the dative object. Notice that dative objects are immune to function changing operations like passive. Nominative agrees in the passive clause with the theme argument *ein Päckchen*. Both sentences reflect unmarked word order. Nevertheless, listeners run into a garden path (indicated by #) in (10a) significantly more often than in (10b). The reason must be that the relative pronoun's Case exerts an influence on the Case-ambiguous nominal *Maria*, which is confirmed in the passive clause in (10b) but contradicted in the active clause (10a): *... *dass der Maria_{DAT} ein Päckchen_{NOM} geschickt hat* could only be rescued if the theme argument *ein Päckchen* were interpreted as an accusative. But this would require a null subject. Since German does not allow null subjects, the parse with

Maria as a dative will invariably fail. (We will return to these studies cited here in Sect. 2.2 when discussing successful processing of lexical Case.)

These examples illustrate that inverse Case attraction is not confined to the correlative/left-dislocation constructions that had been identified in earlier research but occurs in relative clauses with a Case ambiguous head NP in general.

Cinque (2015) refers to data from varieties of Albanian, Farsi and Finnish which show morphologically overt inverse Case attraction as an optional process. In the examples, accusative Case is attracted from a zero nominal in the relative clause to the (nominative) subject of the matrix clause. As far as we can infer from the data, this process does not take place in correlative/left-dislocation constructions either. The same is true of inverse Case attraction in Icelandic as described by Wood et al. (2017). Icelandic speakers appear to prefer sentences in which a dative that originates as a relative clause subject overwrites the nominative Case of the matrix subject. Again, the examples and materials from this study do not involve correlative/left-dislocation constructions.

Considering the syntactic constellation, there are three theoretical lines for the analysis of relative clauses: (i) the Head-External Analysis (HEA), (ii) the Head-Raising Analysis (HRA), and (iii) the Matching Analysis (MA).⁷ The HEA is usually out of focus because relative clauses show connectivity effects which cannot easily be captured if there is no connection but an index between head DP and the trace of the operator inside the relative clause. The HRA assumes that the head-NP is raised out of the relative clause. It explains connectivity effects but has, among other problems,⁸ trouble explaining why the case of the head-NP can differ from the Case that is assigned to the relative pronoun inside the relative clause. The MA is a compromise. It assumes raising of the head-DP to SpecCP of the relative clause plus a copy of this DP in the ‘head’ position. Matrix Case is assigned to the latter, the case of the relative clause to the former.

From the viewpoint of inverse Case attraction, it looks as if the HRA would be the role model.

- (11) Maria [_{CP} [_{der} Maria] C [_{TP} ich [_{der} Maria] begegnet bin]]
Maria who.DAT I encountered am
 ‘Maria, who I ran into ...’

Here it is expected that *Maria* bears dative Case. But this analysis predicts that the head noun *invariably* bears the dative Case that is assigned inside TP. Inverse attraction would thus be the *standard* case, and it would not make much sense to continue calling this ‘attraction’. As we know, however, Case attraction is a fragile process that sometimes applies and sometimes not. In addition to this, no derivational step in (11) has independent motivation; on the contrary. Notice that we are dealing here, as in the experiments to follow, with appositive relative clauses that follow proper names. Although person names can be determined with a D-word

⁷ The issue is so widely known that we can, for reasons of space, refrain from listing the core literature on this topic once again. Useful overviews can be found in Alexiadou et al. (2000), Bianchi (2002a, b) and chapter 1 of Salzmann (2006).

⁸ See Heck (2005).

in the spoken language, splits lead to ungrammaticality: ... *ich [der Maria] begegnet bin => *Maria bin ich [der Maria] begegnet bin*. A straightforward conclusion is that the analysis in (11) must be taken with reservation. Secondly, subextraction from SpecCP is generally impossible. Furthermore, pied piping of the kind *Maria, deren Bruder ich begegnet bin* ‘Maria, whose brother I ran into’ would lead to insurmountable problems. Thus, explaining inverse attraction along the lines of the HRA—see Bianchi (2002a, b)—is afflicted with various problems.

The MA fares better. If the head nominal is a name, as in all our examples, there are no connectivity/reconstruction effects, and thus the MA is more or less indistinguishable from the HEA. There is a Case-marked relative pronoun in SpecCP; this pronoun is coindexed with the head nominal, here a proper name. In modern German, this is an expression lacking morphologically overt Case (if we are allowed to ignore the Saxon *-s* genitive). All we need to assume is that the overt Case of the pronoun is copied onto the Case-underspecified head nominal.⁹

Coindexation between NP and the relative pronoun is likely to be a key factor in the process of Case transfer that we observe in Case attraction, be it progressive or inverse (regressive) attraction. We have seen that in both instances it is a relatively more marked Case that can be transferred to the constituent with a less marked Case. Let us assume with Georgi and Salzmann (2014, 2017) that the head noun is indeed a probe which is in search of an agreeing goal in its locally c-commanded domain. Progressive Case attraction is then a downward operation. In a unified account, the reverse process, inverse Case attraction, would be an upward process and would therefore be incompatible with the mechanics of probe-goal agreement. But this is by no means necessarily so. In a unificational account of probe-goal agreement such as suggested by Pesetsky and Torrego (2007), feature valuation and feature interpretability are disconnected. An interpretable feature can be unvalued, and an uninterpretable feature can be valued. Transferring this proposal to the Case system, one could say that the Case of a Case-ambiguous head noun corresponds to an uninterpretable (or not fully interpretable) feature that is ‘active’ and as such probes the Case-specified relative pronoun whose Case feature is fully interpretable. Agreement between probe and goal leads to an identification of the Case feature. Which Case feature? Of course, the one which is more marked. The underspecified

⁹ Most work on relative clauses circumvents a discussion of appositive relative clauses. De Vries (2006) is an exception. De Vries argues that appositive relatives are special free relatives that are headed by an empty D-head, and that the structure [DP [D° CP]] is the second conjunct of a conjunction phrase (CoP) in which the first conjunct is the name and the second one is an explication or identification of the first. *Annie, who is our manager* is then analyzed as in (i).

(i) [_{CoP} [_{DP} *Annie*] [_{Co} &: [_{DP} [_D ∅] [_{CP} *who* [~~who~~ *is our manager*]]]]]

With respect to inverse attraction, the question would be how the Case of the *wh*-phrase could access the first conjunct. Raising from the second to the first in coordinative constructions has not been observed anywhere else. Raising would have to cross two projections, the one headed by the empty D-head, and the projection of the conjunction head &:. Therefore, it is difficult to see how Case attraction could proceed under these structural conditions.

one is the loser. For this reason, we do not see any serious theoretical problem in extending the essence of Georgi and Salzmann's proposal to inverse Case attraction.¹⁰

Does inverse Case attraction disturb the picture of left-to-right online processing? Certainly not. The parser has a certain window. For Case attraction, be it progressive or regressive, this window needs to cover only two words, the head noun and the relative pronoun (or a zero operator).¹¹

To sum up, inverse Case attraction is widely attested in relative clauses and clearly not limited to special constructions such as left dislocation. In the same way as progressive attraction, inverse attraction is under the control of an asymmetry among the Cases, be it defined according to the familiar Case hierarchy or according to a split between structural and lexical (alias inherent) Cases. Case attraction is in many cases not mandatory and can thus be seen as an effect of linguistic performance, even if this has not received much attention in past research. In the following section, we will look more thoroughly into this largely unexplored area.

2.2 Case attraction and successful processing of Case in language comprehension

Diachronic and synchronic evidence from Case attraction in line with the literature cited above can be seen as evidence of Case attraction in sentence production. Case attraction in comprehension, however, is usually not recorded diachronically, but studied in psycholinguistic experiments.

From the psycholinguistic perspective, Case attraction in comprehension reflects a failure of the parser to keep track of the Cases assigned, leading to a garden-path when the wrongly remembered Case of an argument clashes with the information provided by the verb.¹² The implicit logic of most stimuli in the psycholinguistic literature on Case attraction relies on a combination of two effects:

- (i) We expect Object-Subject (OS) structures to elicit strong garden-path effects when compared to Subject-Object (SO) structures.¹³

¹⁰ As a matter of fact, Pesetsky and Torrego (2007) use examples from Case agreement in Latin as an empirical motivation of their account.

¹¹ Recall that Case attraction takes place under adjacency. Extraposed relative clauses are excluded from Case interactions with the head nominal. Regressive processes are widely known from phonology, the classical domain of strictly local processes. One example would be regressive place assimilation as in German *un+bequem* 'uncomfortable' \Rightarrow *u[m]bequem*, another one Umlaut as in *rot* 'red' \Rightarrow *röt+lich* 'reddish', i.e., essentially a regressive process of vowel harmony.

¹² For attraction phenomena in general, we follow the Hybrid Account of attraction proposed by Häussler (2009). We will discuss the relation of Case attraction to other attraction phenomena and the timepoint at which it occurs in more detail in the General Discussion.

¹³ All German verbs (with the exception of a subgroup of NOM-DAT verbs) are assumed to have unmarked SO structures. Bader and Meng (1999: 122) write: "Evidence from Dutch (Frazier 1987) as well as from German (...) (Bader 1996; Hemforth et al. 1993; Schriefers et al. 1995) suggests that, at least for some types of subject-object ambiguities, there is a preference for the reading with the subject preceding the object (SO-order). Consequently, readers are garden-pathed when a subject-object ambiguity is resolved in favor of the unpreferred object-before-subject reading (OS-order), [...]"

- (ii) We also expect Case attraction to cause garden-path effects. Those occur in structures with embedded relative clauses, where the first argument NP of a matrix clause is directly followed by a relative clause starting with a relative pronoun coindexed with this NP. The first argument NP (CASE-1) can thus attract object Case (CASE-2) from the relative pronoun, and its true Case would be overwritten with object Case. The parser then continues making its way through the sentence, assuming that the word order is OS (marked by the attracted object Case). Once the word order of the matrix clause is revealed on the final verb, there will be a strong garden-path when the structure turns out to be SO instead of the wrongly assumed OS.

Those two garden-path effects can be combined and cancel each other out: If word order in the matrix clause is OS and Case attraction occurs, the OS-garden-path can be attenuated by Case attraction. Because of Case attraction, the parser assumes the structure is OS. When the structure turns out to really be OS instead of the usually preferred SO once the sentence-final verb provides this information, the parser does not run into the OS garden-path, because it was expecting the OS structure since the point where Case attraction occurred. Thus, although Case attraction is a mistake pushing the interpretation into the direction of a dispreferred word order, it can also attenuate the garden-path effect arising when this dispreferred OS word order happens to be confirmed on the sentence-final verb.

Different studies have shown attenuation of OS garden-paths via Case attraction. Using speeded grammaticality judgements of sentences with unambiguously marked Case on the relative pronouns, Bader and Meng (1999) found that DAT can overwrite NOM, while ACC cannot overwrite NOM (at least, not to the extent that DAT does). NOM does not overwrite either one of the object Cases. Bader et al. (2000: 55–63) argue that Case attraction shows parallels to number attraction, based on their observation that marked features (plural, dative Case) can overwrite unmarked features (singular, nominative Case), but not vice versa; see, however, Schlesewsky et al. (2001) for a different interpretation of the data on Case available at that time.

In a set of studies using speeded grammaticality judgements and a mix of morphologically unambiguous and ambiguous Case marking on relative pronouns, Bader and Bayer (2006: ch. 5) again find effects of the attraction of lexical object Case (DAT), but not of structural object Case (ACC). In their first experiment, the authors report significantly more false rejections for SO structures with embedded dative relative clauses (see 12b) than for SO structures with embedded accusative relative clauses (see 12a).

(12) a. Relative clause with structural Case

Daß Anita, die ich übrigens nächste Woche besuchen
that A. who.ACC I by.the.way next week visit
 werde, die neuen Bücher sehr schnell geliefert hat.
will the new books very quickly delivered has
 ‘... that Anita, who I will visit next week, delivered the new books
 very quickly.’

b. Relative clause with dative Case

daß Anita, der ich übrigens letzte Woche begegnet bin,
that A. who.DAT I by.the.way last week met am
 die neuen Bücher sehr schnell geliefert hat.
the new books very quickly delivered has
 ‘... that Anita, who I met last week, delivered the new books
 very quickly.’

According to the authors, this suggests that in (12b), dative Case was attracted from the relative pronoun onto the head nominal, causing an OS interpretation that led to a garden-path effect when the SO structure was revealed on the sentence-final auxiliary. In (12a), accusative Case was not attracted, leading to no processing difficulty for the standard SO structure.

For OS structures with dative objects as in (13), the difference between accusative and dative relative clauses did not cause garden-path effects—all sentences were judged correctly around 80% of the time, just like sentences without embedded relative clauses.

(13) a. Relative clause with structural Case

Daß Anita, die ich übrigens nächste Woche besuchen
that A. who.ACC I by.the.way next week visit
 werde, die neuen Bücher sehr schnell geliefert wurden.
will the new books very quickly delivered were
 ‘... that the new books were delivered to Anita, who I will visit
 next week, very quickly.’

b. Relative clause with dative Case

daß Anita, der ich übrigens letzte Woche begegnet
that A. who.DAT I by.the.way last week met
 bin, die neuen Bücher sehr schnell geliefert wurden.
am the new books very quickly delivered were
 ‘... that the new books were delivered to Anita, who I met last week,
 very quickly.’

The authors interpret their findings as reflecting Case attraction of dative but not of accusative Case. They conclude that Case attraction should not be explained along the Case hierarchy, i.e., the general difference between more and less marked Cases, but instead along the difference between lexical and structural Case. They further conclude that lexical Case may overwrite structural Case, but not vice versa, while structural Cases may not overwrite each other at all.

The studies we present in this paper are intended to build on and extend these earlier studies on Case attraction. Specifically, we are interested in the potential influence of Case attraction on online and offline measures of sentence comprehension. Online measures of sentence comprehension like speeded judgements or self-paced reading times should show signs of Case attraction, in line with the findings in the literature outlined above. In addition, self-paced reading time measurements should enable us to learn more about the time-course of Case attraction in sentence comprehension, how fast it arises and how pervasive and long-lasting it is. One could argue that offline measures like non-speeded judgements may differ by showing no or only weakened effects of attraction.

Online measures of sentence comprehension are influenced by a number of different factors, some of which might interact with or override the OS garden-path—the attenuation of which is the key diagnostic for Case attraction in comprehension. Therefore, the studies presented in the current article use stimuli with structures as parallel as possible in all respects but Case marking: Unlike in the previous studies, all conditions have identical voice and animacy patterns, all conditions to be directly compared are of equal ambiguity/non-ambiguity, and Case marking is provided by two-place verbs assigning either NOM-ACC or NOM-DAT. (The stimulus design will be outlined in detail in Sect. 3.)

The exclusive use of two-place NOM-DAT verbs as dative Case assigners¹⁴ enables us to control for a number of potentially confounding factors relevant for online comprehension measures in our stimuli. However, it introduces a potential new factor influencing reading times and judgements, namely, lexical Case marking effects. NOM-DAT verbs differ from NOM-ACC verbs in sentence comprehension in the following ways:

- Lexical Case marking verbs (NOM-DAT verbs) cause an enhanced processing load (Bader et al. 2000a, b).
- Word order is processed differently (Bader et al. 1996, Hopf et al. 1998, 2003; Bornkessel et al. 2004).
- The animacy of argument referents is processed differently (Czypionka 2014; Czypionka et al. 2017).

These processes reflect *successful* comprehension of NOM-DAT verbs instead of Case attraction. They are known to influence online measures of sentence comprehension. In general, processing NOM-DAT verbs leads to enhanced processing cost in SO-order while it leads to a decrease in processing cost in OS-order.

In the following, we will give a short overview of the effects of successful processing of NOM-DAT verbs, before returning to their relevance in the current context of inverse Case attraction.

¹⁴ In contrast to other strategies like, e.g., different passive choices in three-place verbs, see, e.g., Bader and Bayer (2006), Experiment 1.

In general, NOM-DAT verbs have been shown to have increased processing costs in comparison with NOM-ACC verbs when morphological Case marking on the arguments is ambiguous. Using speeded grammaticality judgements, Bader et al. (2000) found garden-path effects for verb-final sentences ending in NOM-DAT instead of NOM-ACC verbs. The increased workload caused by a reassignment of dative instead of accusative has also been shown to influence ERP measurements. In German verb-final sentences with object-subject-verb word order and with a relative clause between object and subject, Hopf et al. (1998) found an N400, but no P600, with NOM-DAT compared to NOM-ACC verbs. Hopf et al. (2003) were able to replicate this finding in a new experiment. In the stimuli of both studies, the verb of the relative clause assigned accusative, while the final verb in the matrix clause assigned either dative or accusative. The authors interpret the lack of a P600 for the reassignment of dative instead of accusative as a sign that the increased workload of NOM-DAT compared to NOM-ACC verbs is caused by reaccessing the lexical entry of the object, rather than simply by the increased syntactic complexity of dative verbs.

Bader (1996) showed that while OS sentences cause strong garden-paths with accusatives, there is no comparable OS-garden-path with datives. In line with those earlier behavioral findings, an ERP study (Bornkessel et al. 2004) found that OS word orders caused a P600 when compared to SO word orders for NOM-ACC verbs, but N400 without P600 for NOM-DAT. These results suggest that OS word orders cause different types of processing challenges for the two verb classes, and that they do not cause the same extent of syntactic reanalysis for NOM-DAT as for NOM-ACC verbs. Furthermore, the effects resulting from the increased processing cost of animate–animate argument sequences when compared to animate-inanimate argument sequences are weaker for NOM-DAT verbs than for NOM-ACC verbs (Czypionka 2014; Czypionka et al. 2017).

In sum, sentences with NOM-DAT verbs are processed differently from NOM-ACC verbs, and argument linking strategies seem to be different for both, leading to different processing of OS compared to SO structures. When looking for the effects of Case attraction, however, we will manipulate exactly this combination of factors—different word orders with lexical versus structural Case, one assigned by NOM-DAT verbs, the other by NOM-ACC verbs. Given different word order effects for NOM-ACC and NOM-DAT verbs, how can we be sure that we are measuring Case attraction (a reflection of a processing mistake) and not general effects of lexical Case marking (a reflection of ultimately successful in-depth parsing processes)? In the experiments reported here, we set out to find effects of Case attraction that cannot be reduced to general effects of lexical Case marking.

Bearing the distinction between both types of effects in mind, we aim to answer the following questions in the experimental part of this article:

- Can we find Case attraction in online measurements of comprehension?
- What is the time course of Case attraction in comprehension?
- Can we disentangle Case attraction from other Case marking effects resulting from successful processing of Case? (Namely, from the fact that NOM-DAT and NOM-ACC verbs seem to cause different reactions under OS word order.)

If Case attraction really plays a role in sentence comprehension, it should cause measurable effects (i.e., attenuation of OS garden-paths because the original word

order has erroneously been interpreted as OS due to Case attraction). This should hold not only in speeded grammaticality judgements, but also in other online measurements of comprehension. Those effects should occur shortly after the potentially disambiguating region in the stimuli.

In the following, we present the results of a series of experiments on Case attraction in sentence comprehension. To disentangle the effects of Case attraction from those of successful processing of lexical Case marking, we prepared two sets of stimuli. The first stimulus set has overt Case marking in a configuration that allows for inverse Case attraction to occur. The second stimulus set differs from the first stimulus set only in the absence of overt unambiguous case marking. The assumption is that the lack of overt unambiguous Case prevents Case attraction, but not the successful processing of NOM-DAT verbs. This difference between the two experiments allows us to distinguish Case attraction effects (expected only with the first stimulus set) from other effects caused by processing differences between sentences with lexical or structural object Case (which should occur with both stimulus sets). The comprehension of both stimulus sets was monitored offline in non-speeded acceptability judgement tasks, and online using self-paced reading time measurements.

3 Language materials

3.1 OVERT stimulus set

We constructed 50 sentence quartets in four different conditions, crossing the conditions [argument] ORDER (SO or OS) and [object] CASE (accusative or dative).

All sentences began with a statement of belief, opinion etc. (*Ich glaube, dass...*, ‘I believe that...’), followed by a verb-final subordinate matrix clause expressing the content of the belief. This matrix clause contained an embedded relative clause after the first argument. An example of a typical sentence quartet is given in Example 1.

The matrix clauses were verb-final sentences with two arguments. The first NP argument was a single male name, followed by a relative clause specifying this first argument ‘Klaus, who...’. The second argument consisted of two conjoined names (‘Ida and Paul’). The verbs were in perfect tense, i.e., a participle followed by an auxiliary. All matrix clause verbs were non-separable active verbs assigning either NOM-ACC or NOM-DAT. Case-differences between NOM, DAT and ACC are not overtly marked on proper names, therefore, information about the verb’s Case marking pattern (NOM-ACC or NOM-DAT) in the matrix clause becomes available only at the participle. The word order of the matrix clause (SO or OS) is revealed at the word to follow, the clause-final auxiliary, via number agreement. After the auxiliary, there was a four-word spillover region with a neutral sentence context (e.g., ‘when he went for a walk’, see Stimulus Example 1).

Stimulus Example 1:SO-ACC: subject-object, accusative

Ich glaube, dass Klaus, den die Leute gerne einladen,
I believe that Klaus.(NOM) who.ACC the people.(NOM) gladly invite.PL
 Ida und Paul getroffen hat, als er spazieren war.
[Ida and Paul].(ACC) met AUX.SG when he walking was.
 ‘I believe that Klaus—who people like to invite—met Ida and Paul when he went for a walk.’

OS-ACC: object-subject, accusative

Ich glaube, dass Klaus, den die Leute gerne einladen,
I believe that Klaus.(ACC) who.ACC the people.(NOM) gladly invite.PL
 Ida und Paul getroffen haben, als sie spazieren waren.
[Ida and Paul].(NOM) met AUX.PL when they walking were.
 ‘I believe that Ida and Paul met Klaus—who people like to invite—when they went for a walk.’

SO-DAT: subject-object, dative

Ich glaube, dass Klaus, dem die Leute gerne zuhören,
I believe that Klaus.(NOM) who.DAT the people.(NOM) gladly listen.to.PL
 Ida und Paul gefolgt ist, als er spazieren war.
[Ida and Paul].(DAT) followed AUX.SG when he walking was.
 ‘I believe that Klaus—who people like to listen to—followed Ida and Paul when he went for a walk.’

OS-DAT: object-subject, dative

Ich glaube, dass Klaus, dem die Leute gerne zuhören,
I believe that Klaus.(DAT) who.DAT the people.(NOM) gladly listen.to.PL
 Ida und Paul gefolgt sind, als sie spazieren waren.
[Ida and Paul].(NOM) followed AUX.PL when they walking were.
 ‘I believe that Ida and Paul followed Klaus—who people like to listen to—when they went for a walk.’

Example 1: Example of a stimulus quartet from the OVERT stimulus set, used in Experiments 1 and 3. Case attraction to *Klaus* is possible. Labels for morphologically ambiguous Cases in brackets.

The embedded relative clauses were verb-final sentences with two arguments that always occurred in OS word order. The clause-initial objects were masculine relative pronouns referring to the first argument of the matrix clause (i.e., the directly preceding NP). Subjects were bare plural NPs, referring to humans (people, grandparents, policemen etc.). Relative clause verbs were active particle verbs assigning the same Case as the matrix clause verb. The Case of the clause-initial object was unambiguously marked on the relative pronoun (*den* for accusative, *dem* for dative). Thus, argument linking in the embedded relative clause was resolved via Case marking on the relative pronoun and via number agreement on the verb.

For the embedded clauses, we used 39 NOM-ACC and 30 NOM-DAT particle verbs. For the matrix clauses, we used 36 NOM-ACC and 23 NOM-DAT simple verbs. We used some verbs for multiple sentences to construct the final total stimulus number of 50 sentence quartets. (Using some verbs more than once was necessary because of the limited number of NOM-DAT verbs in German.) The accusative and dative-assigning verbs in the embedded and the matrix clause were matched for length and frequency according to the dlexDB corpus (Heister et al. 2011). (Embedded clause: mean length NOM-ACC = 9.1, s.d. = 1.6, mean length NOM-DAT = 9.9, s.d. = 2.2; length $t(49) = -1.98$, $p > 0.05$. Mean frequency NOM-ACC = 1893, s.d. = 2777, mean frequency NOM-DAT = 662, s.d. = 749, t.test log. frequency: $t(49) = 0.67$, $p > 0.5$, frequency unavailable for one embedded NOM-ACC verb, namely *ausbuhen* ‘to boo’. Matrix clause: mean length NOM-ACC = 8.8, s.d. = 1.7, mean length NOM-DAT = 8.8, s.d. = 1.7; length $t(49) = -0.11$, $p > 0.9$; mean frequency NOM-ACC = 13,779, s.d. = 40,311, mean frequency NOM-DAT = 5114, s.d. = 6444; frequency: $t(49) = -0.62$, $p > 0.5$; frequency unavailable for one matrix clause NOM-DAT verb, namely *zürnen* ‘to be angry (with someone)’.) All arguments in the matrix clause and the embedded clause were animate.

The stimuli outlined in the preceding subsection were used in two experiments, an acceptability judgement study (Experiments 1 and 2, see Sect. 4.1) and a self-paced reading time study (Experiments 3 and 4, see Sect. 4.2). Because of our stimulus design, any differences between accusative and dative conditions found in these experiments are in principle attributable to either Case attraction or successful processing of lexical Case marking.

3.2 COVERT stimulus set

The language materials for the second experiment were exactly parallel to the first experiment (identical argument NPs, verbs etc.), with one important change in the relative clause: We replaced the relative pronoun *den* or *dem* with a genitive-marked DP (*dessen Mutter* ‘whose mother’). Notice that *dessen Mutter* remains without any morphological alternations across the Cases nominative, accusative and dative. This configuration excludes Case attraction from the embedded clause to the first NP of the matrix clause. The argument linking pattern in the embedded clause was resolved via number agreement on the verb. All sentences are grammatical. An example of the stimuli used in the second set of experiments is given in Stimulus Example 2.

Stimulus Example 2:SO-ACC: subject-object, accusative

Ich glaube, dass Klaus, dessen Mutter die Leute gerne einladen,
I believe that Klaus.(NOM) whose mother.(ACC) the people.(NOM) gladly invite.PL
 Ida und Paul getroffen hat, als er spazieren war.
[Ida and Paul].(ACC) met AUX.SG when he walking was.
 ‘I believe that Klaus—whose mother people like to invite—met Ida and Paul when he went for a walk.’

OS-ACC: object-subject, accusative

Ich glaube, dass Klaus, dessen Mutter die Leute gerne einladen,
I believe that Klaus.(ACC) whose mother.(ACC) the people.(NOM) gladly invite.PL
 Ida und Paul getroffen haben, als sie spazieren waren.
[Ida and Paul].(NOM) met AUX.PL when they walking were.
 ‘I believe that Ida and Paul met Klaus—whose mother people like to invite—when they went for a walk.’

SO-DAT: subject-object, dative

Ich glaube, dass Klaus, dessen Mutter die Leute gerne zuhören,
I believe that Klaus.(NOM) whose mother.(DAT) the people.(NOM) gladly listen.to.PL
 Ida und Paul gefolgt ist, als er spazieren war.
[Ida and Paul].(DAT) followed AUX.SG when he walking was.
 ‘I believe that Klaus—whose mother people like to listen to—followed Ida and Paul when he went for a walk.’

OS-DAT: object-subject, dative

Ich glaube, dass Klaus, dessen Mutter die Leute gerne zuhören,
I believe that Klaus.(DAT) whose mother.(DAT) the people.(NOM) gladly listen.to.PL
 Ida und Paul gefolgt sind, als sie spazieren waren.
[Ida and Paul].(NOM) followed AUX.PL when they walking were.
 ‘I believe that Ida and Paul followed Klaus—whose mother people like to listen to—when they went for a walk.’

Example 2: Example of a stimulus quartet from the COVERT stimulus set, used in Experiments 2 and 4. Case attraction to *Klaus* is impossible. Labels for morphologically ambiguous Cases in brackets.

4 Offline measurements: acceptability judgements

The stimuli described in the preceding section were tested in two separate acceptability rating experiments. In contrast to earlier studies on Case attraction, we used non-speeded Magnitude Estimation (see Bader 2012). This measurement allows for more fine-grained distinctions between the ratings of different conditions than simple binary decisions. In addition to providing a general overview of the relative acceptability of the different stimulus conditions, it allows us to assess which factors (word order, case marking) affect in-depth parsing processes even without time constraints.

Assuming that Case attraction is very much a processing phenomenon, we expect signs of attraction to be weakened under conditions of metalinguistic judgements. We have no special expectations about the difference between the two Cases in question; however, we expect OS structures to be less acceptable than SO structures in both stimulus sets.

4.1 Methods

4.1.1 Experiment 1: Acceptability judgements, overt Case

Language Materials: The language materials used in the first experiment were the OVERT stimulus set, described in Sect. 3.1. An example is given in Stimulus Example 1.

Participants: 50 participants were tested. All participants spoke German as their only native language. They had no known neurological or reading-related problems. Two participants were excluded from the analysis because they reported being confused by the rating scale. The remaining 48 participants were between 19 and 35 years of age. The mean age was 23 (s.d. = 3.2). Ten of the participants were male. Participants received 2 Euros compensation.

Procedure: The experiment was run in the Psycholinguistics Lab of Constance University. All sentences were rated relative to a reference sentence. The reference sentence was *Die Mitarbeiter haben dass der Chef Probleme hat wohl nicht sofort bemerkt*. ‘Apparently, the coworkers did not notice right away that the boss was having problems.’. It is a grammatical albeit highly marked construction in which the *dass*-CP has been scrambled to the left of the TP. The acceptability for this reference sentence was set to 50. Participants were instructed to rate sentences with higher acceptability with higher values, and sentences with lower acceptability with lower values. The lower limit for (bad) ratings was 1; there was no upper limit to the possible ratings. Before the start of the actual experiments, participants rated five practice sentences. All participants saw 50 critical sentences (12 or 13 per condition), interspersed with 80 filler sentences. Stimuli were presented and ratings were recorded in Linger (Rohde 2003), on the same computer detailed below for the self-paced reading time studies.

Data preparation: The worst value that could possibly be assigned to a sentence was 1. On nine occasions, participants rated sentences with ‘0’. We assumed those

sentences were rated as unacceptable, and manually corrected them to ‘1’ for inclusion in the final data analysis. Following Bader (2012), the rating values were normalized by dividing each rating by 50 (the reference value) and subsequent logarithmizing. Outliers were defined as values that deviated more than two standard deviations from a participant’s mean per condition, and were removed before the final data analysis. 5.4% of the raw data judgements were removed as outliers. Data were prepared for statistical analysis in R (R Development Core Team 2005), using core functions and the packages reshape (Wickham 2007), plyr (Wickham 2011) and car (Fox and Weisberg 2011).

4.1.2 Experiment 2: Acceptability judgements, covert Case

Language Materials: The language materials used in the second experiment were the COVERT stimulus set, described in Sect. 3.2. An example is given in Stimulus Example 2.

Participants: 50 participants were tested. All participants spoke German as their only native language. They had no known neurological or reading-related problems. One participant was excluded from the analysis because she reported having misunderstood the task. The remaining 49 participants were between 19 and 32 years of age. The mean age was 23 (s.d. = 2.9). 12 of the participants were male. Participants received 2 Euros compensation.

Procedure: The procedure, equipment and data preparation were the same as described for Experiment 1. 4.9% of the raw data were removed as outliers before the final data analysis.

4.2 Results of acceptability judgement experiments

The mean of normalized ratings over participants is given in Table 1 for Experiment 1 (overt Case), and in Table 2 for Experiment 2 (covert Case). Graphs were prepared using the ggplot2 package (Wickham 2009) in R.

Ratings were analyzed using a linear mixed-effects model in R, using the packages lme4 (Bates et al. 2015, lmer function) and LMERConvenienceFunctions (Tremblay and Ransijn 2015, summary function). For this first model, we defined the main effects and interactions of CASE, ORDER and EXPERIMENT as fixed effects, and participant and item as random effects. In addition, the main effects and

Table 1 Acceptability ratings for Experiment 1 (OVERT stimulus set), mean normalized acceptability ratings over participants, standard error in parentheses

Condition	Ratings
SO-accusative	0.23 (0.05)
SO-dative	0.19 (0.05)
OS-accusative	-1.13 (0.02)
OS-dative	-1.07 (0.02)

Table 2 Acceptability ratings for Experiment 2 (COVERT stimulus set), mean normalized acceptability ratings over participants, standard error in parentheses

Condition	Ratings
SO-accusative	0.09 (0.06)
SO-dative	-0.08 (0.07)
OS-accusative	-1.59 (0.22)
OS-dative	-1.58 (0.21)

interaction of CASE and ORDER were specified as random slopes for each participant to avoid spurious significances (Barr et al. 2013).

There was a statistically significant main effect of ORDER ($t = 7.3, p < 0.001$), and a statistically significant interaction of CASE and ORDER ($t = -2.3, p < 0.05$). A full table of the statistical results for this first model is given in Appendix A in Table 5.

To pursue the interaction between CASE and ORDER, we ran two additional analyses.

To assess the main effect of ORDER, we separated the data into ACCUSATIVE and DATIVE conditions. The separate datasets were analyzed with a second linear mixed-effects model, with ORDER as a fixed effect, and participant and item as random effects. The main effect of ORDER was specified as a random slope for each participant. A full table of the statistical results for this second model is given in Appendix A in Table 6. For both accusative and dative conditions, the main effect of ORDER was statistically significant (accusative: $t = 9.3, p < 0.001$; dative: $t = 9.3, p < 0.001$).

To assess the main effect of CASE, we separated the data into SO and OS conditions. The separate datasets were analyzed with a third linear mixed-effects model, with CASE as a fixed effect, and participant and item as random effects. The main effect of CASE was specified as a random slope for each participant. A full table of the statistical results for this second model is given in Appendix A in Table 7. For SO conditions, the main effect of CASE was statistically significant ($t = -3.6, p < 0.001$). For OS conditions, there was no statistically significant main effect of CASE.

4.3 Discussion of acceptability judgement experiments

The results of both acceptability judgement studies support the existence of a garden-path for OS compared to SO structures. OS structures are equally dispreferred for accusative and dative conditions, showing no signs of an enhanced acceptability of OS structures with the dative verbs used in our stimuli. This strong dispreference of OS compared to SO structures for both accusative and dative verbs seems to hold across overt and covert Case marking.¹⁵

¹⁵ Given that the sentences were presented without a context that motivates the use of OS-order, this result is not surprising at all.

For SO structures, the results of our acceptability judgement studies reveal a small dispreference for SO-dative compared to SO-accusative structures. Again, this dispreference seems to hold for both overt and covert Case marking stimuli. SO sentences with NOM-ACC verbs were rated as more acceptable than SO structures with NOM-DAT verbs, although SO structures with a DAT-object were not rated as badly as each of the OS structures. While SO structures with an ACC-object received higher ratings than the reference sentence, SO structures with a DAT-object received ratings close to those for the reference sentence, indicating that they were nevertheless judged as relatively acceptable.

We tentatively interpret the dispreference of SO-dative compared to SO-accusative as a reflection of the slightly higher processing load of dative compared to accusative structures, possibly due to a more complex structure for datives than for accusatives. This possibility will be taken up again in the general discussion.

5 Online measurements: self-paced reading times

In a second set of experiments, we monitored the comprehension of the OVERT and the COVERT stimulus set using self-paced reading time measurements. Unlike non-speeded acceptability judgements, but just like speeded acceptability judgements, self-paced reading times are susceptible to errors resulting from processing under time constraints. Thus, self-paced reading could provide a window into the time course of Case attraction if it plays a role in online comprehension.

Contrasting between the OVERT and the COVERT stimulus set allows us to distinguish between effects of Case attraction and effects related to the successful processing of NOM-DAT verbs versus NOM-ACC verbs.

For the first self-paced reading time study (Experiment 3), the OVERT stimulus set was used. For the second self-paced reading time study (Experiment 4), the COVERT stimulus set was used (see Sect. 3 for a detailed outline of the stimulus materials).

In general, OS sentences are harder to process than SO sentences. Therefore, we expect SO-accusative sentences to be processed faster than OS-accusative sentences. We assume that this processing difficulty reflects the extra workload for restructuring the syntactic representation from SO to OS in the matrix clause by reassigning accusative instead of the default nominative Case to the first NP (the proper name *Klaus*).

Assuming that Case attraction does play a role in sentence comprehension, we make the following predictions for the OVERT stimulus set (used in Experiment 3):

- We expect slower reading times in OS structures than in SO structures as soon as the word order information for the matrix clause has become accessible (i.e., from the auxiliary onward).

- If DAT can overwrite NOM, the difference between SO and OS structures should be smaller for dative conditions than for accusative conditions (either due to increased acceptance of object-initial structures, or due to increased spurious rejection of subject-initial structures).¹⁶
- Irrespective of whether ACC cannot overwrite NOM at all, or whether it does so less often than the more marked DAT (see Bader and Bayer 2006 for a discussion), we expect the difference between SO and OS structures to be stronger for accusative than for dative conditions.

For the COVERT stimulus set (used in Experiment 4), we expect any effects related to the successful comprehension of dative-assigning verbs to be stronger than in the OVERT stimulus set. Case attraction, however, should be impossible with this stimulus set.¹⁷

5.1 Methods

5.1.1 Experiment 3: Self-paced reading times, overt Case

Language Materials: The language materials used in Experiment 3 were the OVERT stimulus set, described in Sect. 3.1. An example is given in Stimulus Example 1.

Participants: 54 participants were recruited via flyers and the LingLabs participant mailing list. All participants spoke German as their only native language. Two participants were removed from the data set before analysis because they had given wrong answers to five or more of the ten questions concerning the critical sentences. The remaining participants reported no known neurological or reading related problems and had normal or corrected to normal vision. They were between 18 and 34 years old (mean age 22.02 years, s.d. = 3.17). Thirty-eight participants were female. Participants received 6 Euros compensation.

Procedure: Sentences were presented in a word-by-word, non-cumulative self-paced reading paradigm. Stimuli were presented on a 17" cathode ray tube monitor (Sony Trinitron Multiscan G400), connected to a Fujitsu personal computer. Response latencies were recorded via a keypress on a Razor Deathstalker essential gaming keyboard with a 1000 Hz ultrapolling rate. Stimuli were presented and reaction times were measured using Linger (Rohde 2003). During the experimental session, each participant saw 202 sentences. Fifty of them were critical sentences, the remaining 152 sentences were filler sentences. Each participant saw one sentence of each critical sentence quartet, resulting in either 13 or 12 sentences of each of the four conditions. Comprehension questions were asked after 10 critical

¹⁶ Note that the OVERT stimulus set does not allow us to distinguish between the effects of Case attraction and effects of lexical case-marking verbs. We will deal with this distinction in Experiment 4, see Sect. 4.2.

¹⁷ We do not offer predictions for reading times at the beginning of the relative clauses. This is for two reasons: (i) Unlike later positions, these positions do not allow us to disentangle effects of Case attraction from other effects of lexical Case marking, like different word order preferences. (ii) Following Hopf et al. (2003), we assume that the process of inserting an additional KP projection, the syntactic exponent of dative Case, into the syntactic representation is associated with relatively low processing costs. For the motivation of KP, see Section 6.2 below. This would make it unlikely to find statistically significant effects at this position, even if this process were at the root of Case attraction.

sentences and 38 filler sentences in order to give the participants a task and to keep them motivated. For questions concerning critical sentences, the correct answer was ‘no’ in 5 questions and ‘yes’ in 5 questions. At the beginning of the experimental session, participants were trained with four practice sentences.

Data preparation: Reading times shorter than 202 ms¹⁸ or longer than 4000 ms were removed from the dataset. Outliers were defined as values that deviated more than two standard deviations from a participant’s mean per condition per position, and were removed before the final data analysis. In sum, 4.9% of all data points were removed as outliers.

5.1.2 Experiment 4: Self-paced reading times, covert Case

Participants: 58 participants were recruited via flyers and the Linglabs mailing list. All participants spoke German as their only native language. They had no known neurological or reading-related problems. They were between 18 and 34 years of age. Two participants were excluded before final data analysis because they gave wrong answers to more than 5 of the 10 questions concerning the critical sentences. The mean age of the remaining 56 participants was 22.4 years (s.d. = 3.1). Forty-three participants were female. Participants received 6 Euros compensation.

Procedure: The procedure and equipment were the same as described for Experiment 3 (see Sect. 5.1). The organization of the stimulus lists was identical to the one described in Experiment 1, with 50 critical and 152 filler sentences per participant.

Data analysis: Outlier removal and data analysis were the same as described for Experiment 3. 5.1% of the original data points were removed as outliers.

5.2 Results of self-paced reading time studies

For Experiment 3 (overt Case), the mean reading times over participants per position are given in Table 3, and a graph of the mean reading times over participants is given in Fig. 1. Graphs were made in R using the ggplot2 package (Wickham 2009).

For Experiment 4 (covert Case), the mean reading times over participants per position are given in Table 4, and a graph of the mean reading times over participants is given in Fig. 2.

Logarithmized reading times were compared between experiments and conditions for individual word positions. For positions to the left of the auxiliary, only the information about CASE was available. For the auxiliary and the spillover region, the information about ORDER and CASE was available. Here, we expected an interaction of ORDER and CASE, since the information about argument order only became available with number agreement on the auxiliary.

We focused our analysis on the auxiliary (the first point in the sentence where information about argument order became available), and the four post-auxiliary words in the sentence-final spillover region. Our main expectation was for an effect

¹⁸ 202 ms instead of 200 ms were chosen to exclude a single reading time value of 201 ms.

Table 3 Self-paced reading times for Experiment 3 (OVERT stimulus set), mean reading times over participants in ms, standard error in parentheses

Condition	Position									
	einladen	Ida	und	Paul	getroffen	hat	als	er	spazieren	war
SO-acc	408 (17)	365 (9)	326 (6)	320 (6)	343 (7)	368 (9)	329 (6)	329 (6)	345 (5)	412 (12)
OS-acc	411 (16)	373 (10)	333 (6)	326 (6)	338 (6)	400 (17)	360 (10)	340 (6)	351 (6)	426 (12)
SO-dat	408 (17)	379 (10)	326 (7)	323 (6)	343 (7)	387 (13)	330 (6)	331 (6)	347 (6)	415 (11)
OS-dat	400 (16)	375 (12)	325 (5)	320 (5)	339 (7)	395 (16)	343 (7)	337 (6)	341 (5)	411 (11)

A graph is given in Fig. 1

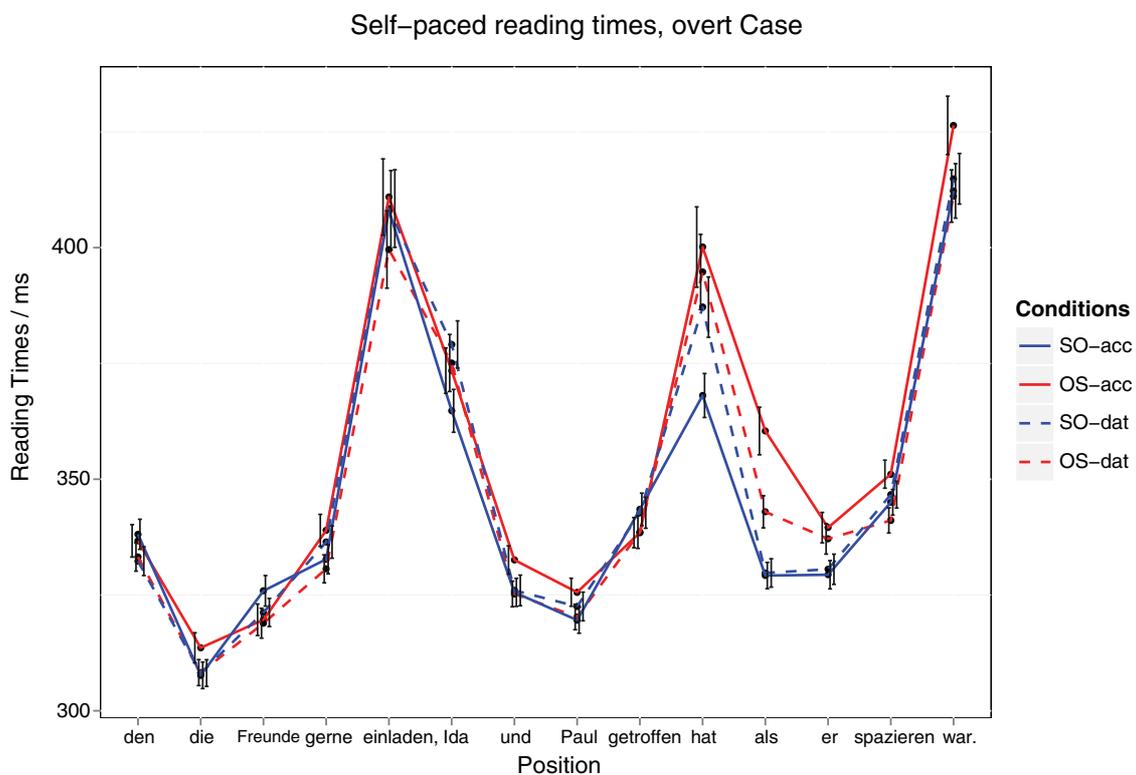


Fig. 1 Self-paced reading times for Experiment 3 in ms, means over participants per condition. Subject-initial conditions are plotted in blue, object-initial conditions are plotted in red. Accusative conditions are plotted as solid lines, dative conditions are plotted as dashed lines. Positions are indicated by the respective words in the SO-accusative condition

of Case attraction on the first post-auxiliary word. Since there is no preceding literature on the time course of Case attraction effects, we offer results for the following positions, too.

Data were analyzed using linear mixed-effects models in R (R Development Core Team 2005), using the packages lme4 (Bates et al. 2015, lmer function) and LMERConvenienceFunctions (Tremblay and Ransijn 2015, summary function). For the comparison of both self-paced experiments, we defined the main effects and interactions of CASE and EXPERIMENT as fixed effects for positions to the left of

Table 4 Self-paced reading times for Experiment 4 (COVERT stimulus set), mean reading times over participants in ms, standard error in parentheses

Condition	Position									
	einladen	Ida	und	Paul	getroffen	hat	als	er	spazieren	war
SO-acc	522 (25)	455 (15)	350 (7)	343 (7)	387 (8)	426 (15)	365 (7)	356 (7)	370 (7)	438 (11)
OS-acc	523 (29)	452 (14)	345 (7)	341 (7)	383 (9)	451 (20)	387 (9)	372 (8)	374 (6)	434 (13)
SO-dat	566 (33)	466 (15)	352 (7)	345 (8)	391 (9)	442 (19)	362 (7)	357 (7)	371 (7)	430 (12)
OS-dat	547 (35)	457 (12)	351 (7)	343 (6)	375 (9)	451 (20)	392 (10)	371 (8)	372 (7)	441 (14)

A graph is given in Fig. 2

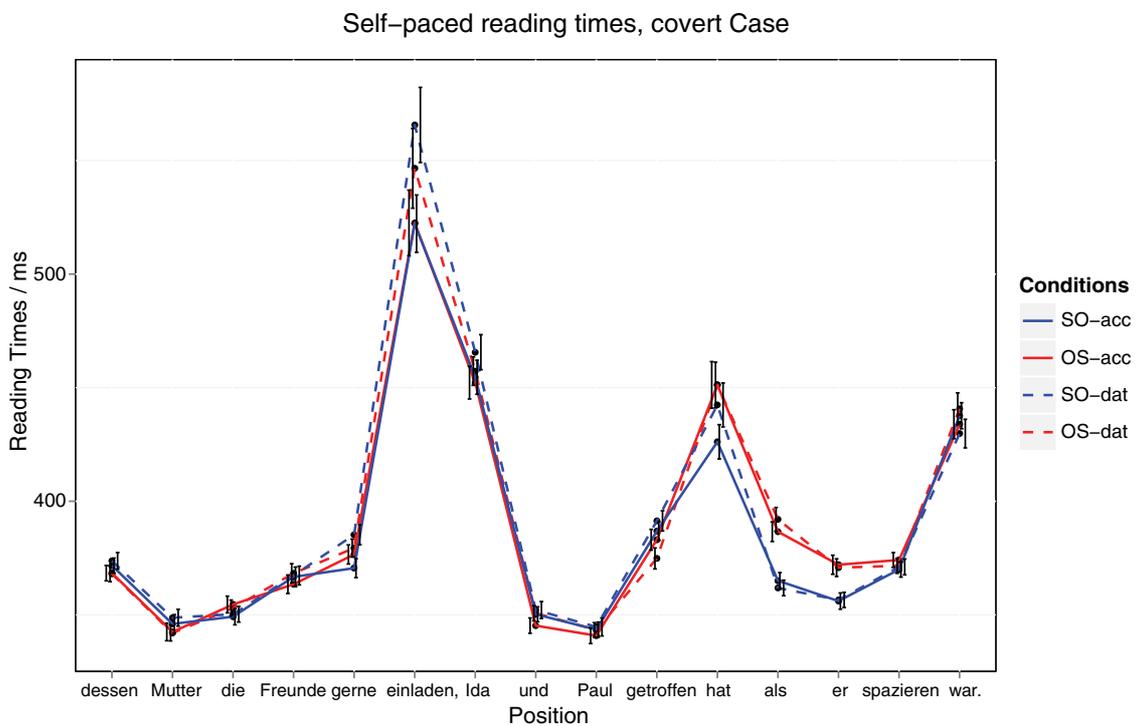


Fig. 2 Self-paced reading times for Experiment 4 in ms, means over participants per condition. Subject-initial conditions are plotted in blue, object-initial conditions are plotted in red. Accusative conditions are plotted as solid lines, dative conditions are plotted as dashed lines. Positions are indicated by the respective words in the SO-accusative condition

the auxiliary. For the auxiliary and post-auxiliary spillover region, the main effects and interactions of CASE, EXPERIMENT and ORDER were defined as fixed effects. Participant and item were defined as random effects. In addition, CASE was specified as random slope for each participant. Effects and interactions are reported and were pursued when they reached or only narrowly missed statistical significance ($p < 0.06$), unless explicitly stated otherwise.

Positions to the left of the auxiliary: No effects of CASE were found to the left of the auxiliary. (Effects of EXPERIMENT are not reported here.)

Auxiliary: There was a marginally significant main effect of EXPERIMENT ($t = -1.9, p < 0.06$). A full table of the statistical results for this position is given in Appendix A in Table 8.

First word post auxiliary (*als*): There was a statistically significant main effect of ORDER ($t = -4.7, p < 0.001$) and a statistically significant main effect of EXPERIMENT ($t = -2.4, p < 0.05$). In addition, there was a marginally significant interaction of CASE, ORDER and EXPERIMENT ($t = 1.9, p < 0.06$). A full table of the statistical results for this position is given in Appendix A in Table 9.

The interaction of CASE, ORDER and EXPERIMENT was pursued in a second model. For this second model, we defined the main effects and interactions of CASE and ORDER as fixed effects, participant and item as random effects, and CASE as random slope for each participant. Data for both experiments were analyzed separately with this model.

For Experiment 3 (OVERT stimulus set), the second model revealed a main effect of CASE ($t = -2.6, p < 0.01$), a main effect of ORDER ($t = -6.1, p < 0.001$), and an interaction of CASE and ORDER ($t = 2.3, p < 0.05$). A full table of the statistical results for this analysis is given in Appendix A in Table 10. To pursue the interaction of CASE and ORDER, we pursued the main effect of ORDER separately for accusative and dative conditions (specifying participant and item as random effects), and the main effect of CASE separately for SO and OS conditions (specifying participant and item as random effects, and CASE as random slope for each participant). The main effect of ORDER was statistically significant for accusative and dative conditions (accusative: $t = -5.9, p < 0.001$, dative: $t = -3.0, p < 0.01$; full table in Appendix A in Table 11). The main effect of CASE was not statistically significant for SO conditions ($p > 0.7$), but was statistically significant for OS conditions ($t = -2.8, p < 0.01$, full table in Appendix A in Table 12). Descriptively speaking, reading times for OS-accusative conditions were longer than for OS-dative conditions.

For Experiment 4 (COVERT stimulus set), the second model revealed a main effect of ORDER ($t = -4.5, p < 0.001$), but no effects of CASE (full table in Appendix A, Table 13). Reading times for OS conditions were longer than for SO conditions.

Second word post auxiliary (*sie*): There was a statistically significant main effect of ORDER ($t = -4.1, p < 0.001$), and a statistically significant main effect of EXPERIMENT ($t = -2.4, p < 0.05$). A full table of the statistical results is given in Appendix A, Table 14. Reading times for OS conditions were longer than for SO conditions.

Third word post auxiliary (*spazieren*): There was a statistically significant main effect of EXPERIMENT ($t = -2.0, p < 0.05$). A full table of the statistical results is given in Appendix A, Table 15.

Fourth word post auxiliary, sentence-final position: There was a statistically significant interaction of ORDER and CASE ($t = -2.2, p < 0.05$) and a statistically significant interaction of CASE, ORDER and EXPERIMENT ($t = 2.6, p < 0.01$). A full table of the statistical results is given in Appendix A, Table 16. The interaction of CASE, ORDER and EXPERIMENT was pursued in a second model. For this second model, we defined the main effects and interactions of CASE and ORDER as main effects, participant and item as random effects, and CASE as random slope for each participant.

For Experiment 3 (OVERT stimulus set), the second model revealed a statistically significant main effect of ORDER ($t = -2.3, p < 0.05$) and an interaction of CASE and ORDER ($t = 2.2, p < 0.05$). A full table of the statistical results is given in Appendix A, Table 17. To pursue the interaction of CASE and ORDER, we pursued the main effect of ORDER separately for accusative and dative conditions (specifying participant and item as random effects), and the main effect of CASE separately for SO and OS conditions (specifying participant and item as random effects, and CASE as random slope for each participant). The main effect of ORDER was statistically significant for accusative conditions ($t = -2.2, p < 0.05$), but not for dative conditions (see Appendix A, Table 18). The main effect of CASE only reached marginal significance for OS conditions ($t = -1.79, p < 0.08$), and was far removed from significance for SO conditions (see Appendix A, Table 19).

Descriptively speaking, reading times in the OS-accusative conditions were longer than in the other three conditions.

For Experiment 4 (COVERT stimulus set), there were no effects on the sentence-final word (see Appendix A, Table 20).

5.3 Discussion of self-paced reading time studies

The self-paced reading time studies revealed main effects of argument order on different positions, the general pattern being that reading times for OS conditions were longer than for SO conditions, after the word order was revealed by the auxiliary.

For Experiment 3 (OVERT stimulus set), but not for Experiment 4 (COVERT stimulus set), this effect of argument order was modulated by the verbal Case marking pattern at two positions. At the position directly following the auxiliary, reading times for the OS-dative condition were shorter than for the OS-accusative condition. On the sentence-final word, the same pattern obtained again for overt, but not for covert Case marking: Reading times for the OS-dative condition were shorter than for the OS-accusative condition.

We interpret this finding as the indication of a reduced OS garden path with overt, but not with covert dative. In the COVERT stimulus set, this garden path is visible in prolonged reading times for OS compared to SO structures, independent of Case marking, descriptively visible on the first two positions following the auxiliary (see Fig. 2). In the OVERT stimulus set, this garden path is visible in prolonged reading times for OS-accusative compared to SO-accusative conditions, descriptively visible on all post-auxiliary positions but the one preceding the sentence-final position (see

Fig. 1). This fits with the assumption that overt dative (but not covert dative, and neither overt nor covert accusative) can be attracted to the head NP, leading the parser to understand the matrix sentence as OS instead of the normally expected SO. This in turn can attenuate the OS garden path effect in overt dative conditions: At least in some instances, OS word order was expected, thanks to Case attraction, and finally confirmed upon encountering an actual OS word order. The effect on the first post-auxiliary word happened on the first spillover region, directly after the position in which information about argument order became available. This suggests that the effects of Case attraction start immediately. The result for the sentence-final position indicates that the consequence can be long-lasting, and would in all likelihood affect later processes of second pass parsing or other revisions.

Importantly, the interaction of CASE and ORDER only holds for sentences with overt Case marking. For covert Case marking conditions, only main effects of ORDER are found. This shows that the interaction of CASE and ORDER cannot be reduced to the potential licensing of flexible word orders as induced by NOM-DAT but not NOM-ACC verbs (see, e.g., Bornkessel et al. 2004). If it were, it should have obtained for both experiments equally. This suggests that any effects of CASE we found in Experiment 3 are effects of morphological Case, and not of either abstract Case nor of other properties related to the two different verb classes involved.

6 General discussion and conclusion

6.1 Interpretation of experiments

In this article, we set out to monitor the time course of Case attraction effects in sentence comprehension, and to disentangle them from the effects of lexical Case marking. To this end, we constructed two stimulus sets that were exactly parallel, with one single difference: While the stimuli in the OVERT set could have elicited both Case attraction and effects resulting from successful comprehension of NOM-DAT verbs, the COVERT stimulus set could have elicited only the latter.

The results of both our offline acceptability judgement studies (Experiments 1 and 2) show a strong dispreference for OS compared to SO conditions, irrespective of verbal Case marking. This OS dispreference was found for both the OVERT and the COVERT stimulus set, suggesting that in offline measurements, the presence or absence of overt Case marking does not influence the relative acceptability of OS structures compared to the respective SO structures. This finding is also important for the interpretation of the self-paced reading experiments. Given the results of the acceptability judgements, the difference found between the two self-paced reading experiments is unlikely to reflect a difference in acceptability between the OVERT and COVERT stimulus set.

For the SO structures, there was a slight dispreference for SO-dative compared to SO-accusative structures. This suggests that processing NOM-DAT verbs leads to a slightly enhanced workload. The reason for this could be a more complex phrase structure that needs to be constructed, a proposal to which we will turn in Sect. 6.2. There was no Case marking effect for OS structures, perhaps because of the strong

dispreference for OS structures that may have led to a bottom effect, obliterating more subtle differences between OS-accusative and OS-dative structures.

The pattern is different in the online measurements of the same stimulus sets, however. In Experiment 3 (using the OVERT stimulus set), we find an attenuation of the OS garden path for the dative but not for the accusative conditions at two different positions—directly following the auxiliary, and at the sentence-final position during sentence wrap-up. In isolation, these effects could be interpreted either as Case attraction or as licensing of OS word order by NOM-DAT verbs (as found in EEG experiments by Bornkessel et al. 2004). However, in Experiment 4 (COVERT Case) only the latter, but not the former, is possible. When our stimuli make Case attraction impossible, we find a robust OS effect, with longer reading times for OS than for SO structures, both in accusative and dative conditions.

We therefore interpret the results of our reading time studies as a reflection of Case attraction in online comprehension, matching earlier speeded grammaticality judgement studies¹⁹ (Bader and Meng 1999; Bader and Bayer 2006). Importantly, the results of Experiment 4 confirm that Experiment 3 indeed reflects Case attraction and no other processes that may also be associated with the comprehension of NOM-DAT verbs.

The difference between the effects of overt and covert Case marking in self-paced reading times could also be attributed to the fact that for the COVERT stimulus set, the embedded sentences were longer by one word and more complicated than for the OVERT stimulus set (because the relative clause began with a phrase rather than a single word), thereby leading to a loss of Case attraction effects due to generally more difficult processing. However, this should not be the main reason for the difference between the two self-paced reading time studies (Experiment 3 and 4). Longer sentences should cause a bigger working memory load, something that has been linked to an increased likelihood of attraction errors (Fayol et al. 1994; Bock et al. 2001: 114). Since it was the shorter sentences of the OVERT stimulus set that elicited Case attraction effects, we are confident that the small difference in length between OVERT and COVERT stimulus sets is not at the root of the different results found for Experiments 3 and 4.

Our findings show an interesting contrast between the results of non-speeded acceptability judgement and reading time tasks. The acceptability judgements did not show any effects of Case attraction, only the strong garden-path caused by OS structures when compared to SO structures, and the mild enhancement of complexity caused by encountering a NOM-DAT configuration. The self-paced reading time studies, in contrast, did show effects of Case attraction (an attenuation of the OS garden-path for NOM-DAT verbs, but only with overt Case morphology). We interpret the self-paced reading time studies as showing the effects of online comprehension. Since inverse Case attraction is a spontaneous transfer between adjacent elements, rather than parsing according to the competence grammar, we assume that measurements of online processing (like self-paced reading times) are more likely to be sensitive to Case attraction than non-speeded judgements like Magnitude Estimations.

¹⁹ The stimuli and methodology in our study differ in too many ways from the ones used in the older studies to allow for a direct comparison between results. However, both our study and the cited studies allow the conclusion that Case attraction affects online measures of sentence comprehension.

How can we integrate our findings into the wider context including also other attraction phenomena? The majority of psycholinguistic studies investigating attraction effects are concerned with number attraction, which does not allow an easy extension to Case attraction. In this literature, a main topic of the current debate is whether (number) attraction is best explained as reflecting feature percolation/transfer, or errors in cue-based retrieval. The most promising account for a tentative discussion seems to be the Hybrid Account proposed for number attraction in Häussler (2009), given that this account in principle allows for both explanations in parallel.²⁰

Häussler points out that the term ‘attraction’ is used to refer to a variety of different processes, reflected in illusions of grammaticality/ungrammaticality in number agreement, but leading to different patterns. In the Hybrid Account of number attraction, Häussler suggests that attraction can occur at two different points during sentence processing. The first one is when the subject NP is encountered, and its number specification is computed. This process is vulnerable to interference from the more marked plural, but not from the unmarked singular. Häussler suggests that parsing errors during this processing step exhibit an asymmetry between marked and unmarked features, and are explained via accounts based on feature transfer (e.g., Bock et al. 2001, 2004; Eberhard et al. 2005), because this type of transfer is only possible with explicitly expressed features. The second timepoint is when the finite verb is encountered and number agreement between subject and verb is computed. This second step is vulnerable to interference from distractors with (partially) matching retrieval cues for subject retrieval. Häussler suggests that parsing errors during this processing step exhibit no asymmetry between marked plural and unmarked singular, thus making feature transfer an unlikely explanation. Cue-based retrieval accounts (e.g., Wagers et al. 2009) are better suited to explain the number attraction errors occurring during this processing step.

Our own findings exhibit an asymmetry between overtly expressed accusative and dative. If we follow the general argumentation of the Hybrid Account (assuming roughly similar processes underlying Case attraction and number attraction), this would suggest that Case attraction is a parsing error that happens at the first time-point proposed by Häussler, i.e., resulting from feature transfer during relatively early processing steps. These early processes (here: the computation of the Case of the first NP) are supposed to be sensitive to markedness contrasts and exhibit an asymmetry. In contrast, if Case attraction were a cue-based retrieval process, happening at a later time-point, we would expect to find no asymmetry, i.e., equal amounts of attraction for accusative and dative.

Given that we find an asymmetry between dative and accusative, we therefore assume some type of feature transfer happening during early processing steps. However, the question remains which features are transferred in the type of Case attraction found in our experiment. Both accusative and dative are explicitly marked in overt morphology on masculine relative pronouns, so the mere presence of overt

²⁰ Häussler (2009) cites Nicol et al. (1997) with the example *The owner of the house who charmed the realtors was no longer willing to sell.* to illustrate number attraction in relative clauses, leading participants to wrongly reject this sentence as ungrammatical.

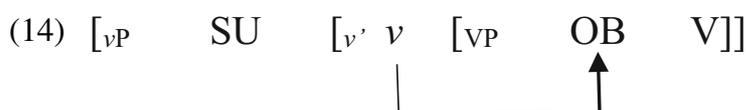
morphological Case marking cannot be the distinguishing factor. Rather, the feature that is attracted must be something that is either categorically present in datives, but not in accusatives, or else much stronger in the former than in the latter. Earlier studies (Bader and Bayer 2006, ch. 5) seem to suggest that the relevant contrast is categorical in nature, i.e., that lexical dative, but not structural accusative, is attracted (however, these earlier studies did not provide online measurements, and did not control for some of the confounds present in the literature). This would suggest that the ‘more marked’ feature in our stimuli corresponds to lexical Case marking instead of structural Case marking. An alternative explanation would be that the relevant contrast between the cases is not categorical, but better explained with a markedness hierarchy, with nominative the least marked, accusative more marked, and dative the most marked of the three Cases employed here. Before Case attraction can be successfully integrated in the wider psycholinguistic literature on other attraction phenomena, future studies will be needed to provide a clear characterization of the nature of the accusative-dative asymmetry in Case attraction.

Importantly, the two possibilities outlined here would not lead to different predictions for Case attraction in our current stimuli; and distinguishing between the two accounts was not the goal of this study. Instead, we aimed to disentangle effects of Case attraction from other effects of lexical Case—for our current issue, finding an asymmetry between overt accusative and overt dative, but none between both covert Cases, is sufficient.

Given this asymmetry, the question is how one can account for it in structural terms and in their consequences for online comprehension. This is the topic of the next section.

6.2 Theoretical interpretation

Before we embark on the interpretation of OS-order, let us first dwell on a plausible expectation about SO-order and how it might be affected by inverse Case attraction. Recall that the auxiliary agrees with the subject. If there is Case attraction from the relativizer to the head nominal, i.e., the proper name, it should be the same for both Cases under consideration. If the parser has developed the hypothesis that the proper name cannot be the nominative subject, this hypothesis has to be revised. The findings of Bader and Meng (1999), Bader et al. (2000), Meng and Bader (2000), and Bader and Bayer (2006) show that this is the case but that the revision process is harder in the case of the dative than in the case of the accusative. As these authors pointed out, the difference can hardly be reduced to the morphological difference between the two forms. To find a better answer we must understand how dative as opposed to accusative case is licensed. Following a standard assumption of Minimalist syntax, the functional head which is responsible for accusative Case is little *v*.



Assume that v has an unvalued feature which is valued as soon as it probes a nominal object. The object may stay inside VP or it may undergo leftward movement for reasons of information structure. This is what we call A-scrambling. Dative Case cannot be licensed along the lines of (14) as v is not responsible for probing the object. If there were a probing relation, there would have to be a functional head different from v .²¹ While there is some evidence that this is possible in certain languages,²² German offers no evidence in this direction. It is widely known that function changing operations such as passive or middle formation do not affect the dative whereas they affect the accusative.²³ In (14), if v is missing, as is the case in passives but also in other unaccusative constructions, there is no way of licensing accusative Case. Datives, however, stay.²⁴ This widely known effect goes hand in hand with a far less widely known effect, namely that accusative does not

²¹ This is exactly what various researchers assume. Woolford (2006), McGinnis (1996) and following work assume that datives with a goal-role are licensed by a separate little v . It will immediately become clear why we do not assume for German a proliferation of functional heads beyond v .

²² Indirect object agreement in Basque may suggest the existence of such a head.

²³ See Haider (1984, 1985), Bayer et al. (2001), Woolford (2006) and many others. An often repeated counter argument is the notorious *bekommen/kriegen* passive as seen in the following:

(i) Wunibald kriegt die Fingernägel geschnitten.
W. gets the finger.nails cut
 ‘Wunibald gets his nails cut.’

(ii) Walburga kriegt ihre Turnhose gebügelt.
W. gets her sport.pant ironed
 ‘Walburga gets her sport pants ironed.’

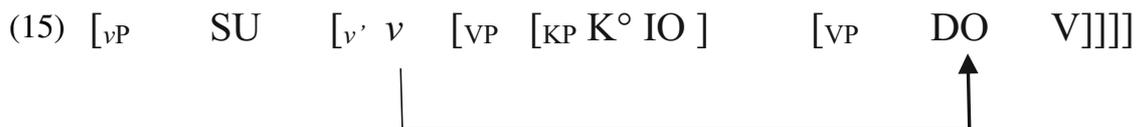
for which it has been argued that dative assignment is suppressed, and the object promotes to the subject position. This kind of ‘passive’ though interesting in its own right should, however, not be confused with the V-dependent passive, a view that is supported by Woolford (2006). Notice also that many simple transitive dative verbs do not allow it at all:

(iii) Willibald bekam von niemandem {*vertraut, *geschadet, *begegnet, *gehört,
W. got by nobody trusted, harmed, encountered, obeyed,
 *nachgerannt, ??zugehört, *beigesprungen, *gegrollt, *widerstanden, *entflohen}
run.after, listened.to assisted rumbled resisted fled

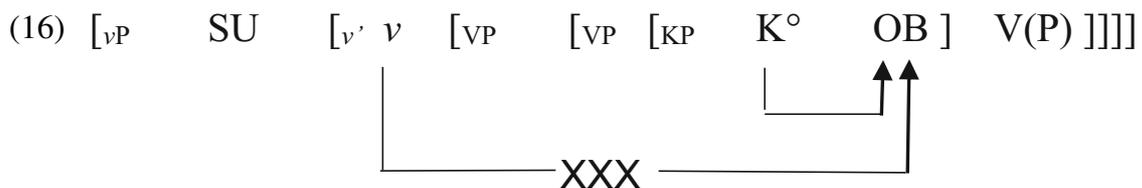
Bader and Häussler (2013), who study mainly passive in ditransitive constructions, come to the conclusion that *bekommen/kriegen* is an emerging passive auxiliary which, however, retains clear traces of the lexical meaning.

²⁴ One reviewer points to potential problems with inherently reflexive verbs like *sich benehmen* ‘to behave oneself’, *sich schämen* ‘to be ashamed’ or Italian *vergognarsi* ‘to be ashamed of oneself’. These are formally transitive but semantically intransitive. Although reflexives can be dative as seen in *Er widerspricht sich* ‘He contradicts himself’, they never seem to be datives or genitives in inherently reflexive verbs that take arguments as in *sich* [_{GEN des Lebens}] *freuen* ‘to enjoy life’ or *sich* [_{DAT dem Klavierspiel}] *widmen* ‘to dedicate oneself to piano-playing’. Assuming that in these cases the reflexive absorbs accusative Case if any, one problem could be the existence of passives like *Jetzt wird sich mal besser benommen!* ‘Behave yourself in a better way!’. Similarly, as Kallulli (2013) points out, deponent verbs like Latin *labor* ‘to slide down’ are passive in form but active in meaning and can occasionally be transitive like Latin *sequor* ‘to follow’. While these are challenging problems of their own, we do not see the immediate relevance of these facts for the distinction between structural versus inherent Case that the reviewer seems to have in mind.

rely on overt Case morphology whereas dative relies very much on it. Bayer et al. (2001) give much evidence in favor of this difference. Their conclusion is that the more or less obligatory Case morphology that we see on datives is not just a morphological wrinkle but must be interpreted as the surface reflex of genuinely syntactic structure. The authors assume that the lack of functional licensing forces the dative argument to establish its own functional structure. According to them, the difference between accusative and dative is not only a featural difference as suggested by the Case hierarchy but a difference in phrase structure. The dative, but not the accusative, is a Kase-phrase, i.e., it involves a Kase-shell that is erected over NP/DP. The K-head does not need to be stipulated (in the sense of an additional little v); K is the functional exponent of the dative's morphology.²⁵ Abstracting away from German and its morphological Case system, K can be seen on a par with the prepositional head that accompanies datives in other languages. English, Dutch, Afrikaans and mainland Scandinavian languages like Swedish use the option of 'P-insertion' in ditransitives: **I gave the book John => I gave the book to John*. The dative is semantically but not morpho-syntactically V-dependent but rather adjoined into the phrase marker as a larger structure. (15) shows the dative in a ditransitive construction where we assume it is adjoined to VP.



(16) below shows the dative as the direct object. As often assumed for the representation of unergatives, i.e., intransitive clauses with an agentive subject role like *John danced*, v may be present but has no job to do for the object.²⁶ Minimality prevents v from probing the dative object because there is a closer functional head that probes the NP/DP. This functional head is, of course, K.



KP is an argument. However, it is not licensed by one of the two functional associates of the finite verb v , and T. Thematic structure is independent of the functional structure. The predicate has a number of theta roles to discharge while the

²⁵ A projection of KP has already been proposed in earlier work by Lamontagne and Travis (1987) and by Bittner and Hale (1996).

²⁶ The only job for v to do is to license the external argument's theta role. We are aware of the questions that emerge from Burzio's generalization as well as the fact that v is in need of a variety of qualifications (see Kallulli 2013), but pursuing these issues would lead us far away from the objective of shedding light on the nature of inverse Case attraction.

syntax of Case may have various ways to map theta roles onto the functional structure of the clause. Nothing in thematic theory requires implementation of the object's grammatical function in terms of a ν - or T-based probing relation. A structural object may, for instance, also be taken care of by incorporation.²⁷ If so, KP is merged like a thematically selected PP, e.g., a goal PP as in *He went to school*, *He ran out of the house*. As Bayer et al. (2001) show, dative arguments behave throughout like selected PPs.

With this much of a minimal proposal about object Case in German, let us now bring Case attraction into the picture again. Given the Matching Analysis (MA), the Case of the relative's DP alias the relative pronoun is independent of the Case of the head-DP (which in the examples under consideration is simply a name). Assume now that the Case of the relativizer is more marked than the Case of the head nominal. All we need to assume is that in this very context the overt Case of the relativizer is—erroneously—copied onto the Case-underspecified head nominal. Given our experimental data, one can be sure that covert/abstract Case as in the relative operator *dessen Mutter* 'whose mother' is not attracted.²⁸

As Meng and Bader (2000), Bayer et al. (2001) and Bader and Bayer (2006) have argued, the difference between the two attractee Cases under debate may be morphologically on a par or close to being on a par, but they are syntactically very different. Transfer of dative Case onto the head NP/DP invokes the construction of a KP-shell over NP/DP while nothing of this sort follows from the copying of accusative Case. Compare (17) with (18).

²⁷ (i) dass er Tag und Nacht {die Waldstein-Sonate
that he day and night the W-sonata
 spielt/ klavierspielt/ *die Waldstein-Sonate klavierspielt}
plays piano.plays the W-sonata piano.plays
 'that he plays {the Waldstein-sonata/the piano} day and night'

(ii) dass er {einen BMW fährt/ autofährt/ *einen BMW autofährt}
that he a BMW drives car.drives a BMW car.drives
 'that he {drives a BMW/is a driver}'

The incorporated direct object detransitivizes the verb by blocking a theta-position. One reviewer remarks that definite objects, which may bear overt Case, may get a so-called 'weak' interpretation analogous to incorporated objects, as in *Hans spielt den Kontrabass* 'Hans plays the double-bass'. There is a clear semantic difference, however. The direct object in the example denotes *the* double-bass that is prescribed in a certain piece of music or that is an essential instrument in a certain ensemble. *Hans spielt Kontrabass* simply means that he is a bass-player or that he is actually playing a double-bass.

²⁸ From the viewpoint of the competence grammar, that naturally lacks a time constraint, covert Case attraction would be a legal option. Assume that the parser arrives at the clause-final verb. Only at this point is it revealed which Case X is required by the verb. The parser would then backtrack to the trace of the relative operator in the A-position and determine Case X. Case X is inherited to the A'-position of the operator. At this point, inverse Case attraction could copy X onto the head nominal. While this is logically viable, it is psycholinguistically unrealistic. Backtracking and reanalyzing the input in second pass parsing is an extreme option that is connected with a heavy workload, and as such it would be completely esoteric in the processes of real-time Case assignment. Our findings as well as those of previous research suggest that inverse Case attraction is an immediate and strictly local process of transfer that may take place as soon as the relative pronoun/operator is received by the parser. This excludes the steps described above, and therefore it excludes the attraction of covert Case.

- (17) $[_{NP} [_{NP} [_{CP} [_{KP} K^{\circ} [_{NP} dem]]] [_{TP} \dots [_{KP} K^{\circ} [_{NP} dem] \dots]]]]]]$
 \Rightarrow COPY DATIVE \Rightarrow
 $[_{KP} [_{KP} K^{\circ} NP [_{CP} [_{KP} K^{\circ} [_{NP} dem] [_{TP} \dots [_{KP} K^{\circ} [_{NP} dem] \dots]]]]]]]]$
- (18) $[_{NP} [_{NP} [_{CP} [_{NP} den] [_{TP} \dots [_{NP} den] \dots]]]]] \Rightarrow$ COPY ACCUSATIVE \Rightarrow
 $[_{NP} [_{NP}_{acc} [_{CP} [_{NP} den] [_{TP} \dots [_{NP} den] \dots]]]]]$

Attraction of the dative triggers an enrichment of the head-NP's phrase structure that is absent under the attraction of accusative. Recall that 'accusative' is a morphological label that serves valuation of the probing head v , and v is part of the functional grid that is by default involved in the transitive and intransitive/unergative clause. The assumption of this difference between structural and 'inherent' Case is very well motivated. Appendix B provides a series of tests according to which the KP-hypothesis for the dative makes the right predictions and therefore motivates a qualitative split between the two structural Cases nominative and accusative and the 'inherent' Cases dative and genitive.

This is the point at which we need to consider how the human parser fares with the two scenarios for inverse Case attraction that the competence grammar offers. The CASE PREFERENCE PRINCIPLES as first formulated in Bader et al. (1996) and then in revised form in Bader and Bayer (2006: 108ff) predict that in the absence of any other (such as morphological) information the first NP in the input will be assigned nominative Case, the second accusative, and the third—as a last resort—dative. Attraction of overt Case interferes with the application of the CASE PREFERENCE PRINCIPLES. Case attraction from a marked pronominal to an unmarked or less marked NP is, as we have argued, a consequence of adjacency, coindexation and the parser's commitment to structural information.²⁹ Under the condition of Case attraction as depicted in (17) and (18), the so far neutral head NP acquires unambiguous Case, dative in (17), accusative in (18). Assume now that the head-NP which up to the respective parsing stages $[_{NP} [_{NP} [_{CP} [_{KP} K^{\circ} [_{NP} dem]]]$ and $[_{NP} [_{NP} [_{CP} [_{NP} den]$ still remains locally ambiguous according to the competence grammar, gets integrated into the ultimate clause structure. The output of Case attraction seen in (17) and (18) respectively will be confirmed by the OSV-scrambling structure with the verb's plural agreement but disconfirmed by the SOV-structure with the verb's singular agreement. Even though scrambling

²⁹ We must remain somewhat vague here because it is not clear which of the three factors is really decisive. Immediate adjacency as such cannot be responsible. If it were, we would expect Case attraction in many cases in which it is not attested. Transfer as in (i) and (ii), i.e., from a dative DP to an adjacent nominative DP, clearly plays no role.

- (i) *als Peter seinem.DAT Chef(.DAT) ...* \Rightarrow *als Peter(.DAT) seinem.DAT Chef(.DAT) ...*
(ii) *als man Peter seinen.ACC Chef(.ACC) ...* \Rightarrow *als man Peter(.ACC) seinen.ACC Chef(.ACC) ...*

The fact that the relative operator picks up the referent of the head NP/DP clearly plays a central role in the Case transfer we see in attraction processes. On the other hand, given the immediacy of Case transfer during parsing, head-NP and relative pronoun must be adjacent. Case attraction into or from an extraposed relative clause has never been observed.

is an unpreferred option, Case attraction enhances the parser's inclination to opt for a scrambling analysis.

This was by and large the result so far. The question is why in Experiment 3 (self-paced reading with unambiguous Case morphology) the SOV condition does not yield the same reversal effect (i.e., why there are no statistically significant penalties for SO-DAT compared to SO-ACC conditions). Even if there is an effect of Case attraction, it does not show the previously established dative effect.³⁰ This difference may be explained by an important difference between the former experiments and the present ones. The former experiments used a passive construction as seen in (10b), *dass Maria, ..., ein Päckchen geschickt wurde* 'that Maria.DAT a parcel.NOM was sent', while the present ones used plain scrambling, as seen in *dass Klaus, ..., Ida und Paul gefolgt sind* 'that Klaus.DAT Ida and Paul.NOM followed'. Notice that scrambling involves a reversal of basic word order, i.e., from SO to OS, whereas this is not true for the passive construction. In the latter, the recipient argument precedes the theme argument as in the active sentence order *dass jemand Maria, ..., ein Päckchen geschickt hat* 'that someone Maria.DAT a parcel.NOM sent'. As a consequence, the OS-condition of the previous experiments should not be confused with the OS-conditions of the present experiments. It is widely known that scrambling is disfavored by the human parser. Thus, the present experiments involve a particularly difficult condition. Our tentative conclusion is that this condition is strong enough to obliterate the more subtle Case difference between the attraction of dative vs. accusative Case.

Consider now the OS-condition of Experiment 3. Here we observe a significant attraction effect in two positions after the auxiliary verb of the relative clause. In particular, an advantage of dative attraction over accusative attraction can be observed at the complementizer *als* and at the final auxiliary of the adverbial clause. In this condition, we compare OS scrambling order, and it turns out that the effect of the purported dative attraction leads to an ameliorization as compared to the the purported accusative attraction.

Important for the present theory is that the head NP is affected by a dative relative pronoun more than it is by an accusative relative pronoun. According to the KP-hypothesis, this means that the first NP may be taken for a dative due to Case attraction. Notice that in the dative condition, according to (17), a piece of phrase structure, namely the KP, has been erected which is confirmed if OSV-order is indeed the order that is revealed by number agreement with the subject (here the plural *Ida und Paul*). The transfer of accusative Case, if it takes place at all, has a weaker effect as it does not change the phrase structure of the input but involves a Case label at best. Thus, attraction is likely to rest on a rather weak effect that does not lead the parser to expect the disfavored OS scrambling construction. In the spillover region immediately after the auxiliary, the morphologically overt dative has a significant advantage over the accusative in the OSV conditions. Case

³⁰ The present study cannot determine a general attraction effect because there was no condition to examine the (non-)influence of the nominative relative pronoun *der*.

attraction can explain the difference between the two verb classes. If the dative has been copied onto the head-NP from the relative pronoun, this NP is enriched by a KP-shell. As soon as the plural auxiliary *haben* is received, the parser has to opt for the generally less accepted OSV order. In the dative condition, the head-NP may already be a potential dative object thanks to dative attraction, and no revision of the current phrase structure is required if the parser has to pursue an OSV-analysis. Why is the OSV-analysis comparatively harder in the accusative condition? The most straightforward explanation would be that accusative Case is not attracted to the head-NP at all; this explanation amounts to the syntax-based account of Case attraction proposed by Bader and Bayer (2006: ch. 5).

Our theoretical interpretation is readily compatible with Häussler's (2009) hybrid account of number attraction. Right after the parser has received the relative (dative) pronoun, there is an optional process of transfer from a marked—in our case KP-headed—pronoun to the head NP, which in this way inherits the source's KP-structure. This process is followed later on by another process that is independent of markedness and simply controls the agreement between subject and verb. If dative attraction has paved the way for the interpretation of the head NP as a dative, the plural agreement between subject and verb confirms an OS scrambling analysis. As in Häussler's account of number attraction, this second process is only indirectly related to Case attraction proper. Plural agreement between the plural subject and the verb, as in the present case, is independent of Case attraction and scrambling. If Case attraction has applied, however, number agreement complies with a scrambling analysis.

At the current stage of our research, we cannot decide between the options of accusative not being attracted at all and thus never overwriting a less marked Case, or accusative attraction as structurally much less prominent than dative attraction. The three confirmed results are that according to the present investigations, (i) there is inverse Case attraction; (ii) there is inverse Case attraction in online comprehension; and (iii) in agreement with earlier results, inverse Case attraction yields much stronger effects for dative than for accusative Case. With respect to (iii), we did not find a convincing dative effect in our offline studies apart from a slight dispreference for dative as compared to accusative Case in the SO-condition. Since this dispreference holds also in the condition with covert Case, the burden of an explanation may lie more on the lexical nature of the dative-assigning verbs than on Case as such.

Where the difference in the prominence of Case attraction cannot be shown experimentally, independent syntactic facts as collected in Appendix B may be considered. If dative Case involves extra phrase structure as suggested by the KP-hypothesis, the difference between dative- and accusative-assigning verbs finds a straightforward explanation, at least in German. Given the highly noticeable structural difference between the two object Cases under debate, it appears to be at least plausible that this difference has repercussions in sentence processing.

In line with earlier studies, our results show that inverse Case attraction in sentence comprehension is psychologically real. The design of our experiments allows us for the first time to completely separate Case attraction effects from the effects of successful processing of NOM-DAT verbs. Both of these play a role in

sentence comprehension, and both processes fit with a KP-based account of dative assignment.

Acknowledgements We gratefully acknowledge clarifying communications with Alex Grosu and Martin Salzmann and the assistance of Christina Gozebina, Marc Meisezahl and Thi Xuan Mai Truongh during data acquisition. The comments by three anonymous reviewers were highly relevant for the improvement of this article. Laura Dörre received funding via a Ph.D. Grant by the *Studienstiftung des Deutschen Volkes*.

Appendix A: Statistical results

R packages used for data preparation, analysis and visualization: lme4 (Bates et al. 2015), LMERConvenienceFunctions (Tremblay and Ransijn 2015), plyr (Wickham 2011), reshape (Wickham 2007), car (Fox and Weisberg 2011).

In Appendix A, we list the summaries of the fixed effects in the mixed-effects models for all experiments. The specific experiments are indicated in each caption. Statistical significances are given as ($p < 0.1$), $*(p < 0.05)$, $** (p < 0.01)$ and $*** (p < 0.001)$.

Statistical results, acceptability judgements

See Tables 5, 6 and 7.

Table 5 Summary of the fixed effects in the first mixed-effects model for Experiments 1 and 2, Magnitude Estimation Studies

	Estimate	SE	t value	Pr(> t)
(Intercept)	-1.596	0.214	-7.46	4.3e-11***
CASE	0.012	0.069	0.17	0.87
ORDER	1.690	0.231	7.31	8.9e-11***
EXPERIMENT	0.463	0.303	1.53	0.13
CASE:ORDER	-0.182	0.078	-2.31	0.03*
CASE:EXP.	0.049	0.098	0.50	0.62
ORDER:EXP.	-0.328	0.327	-1.00	0.32
CASE:ORDER:EXP.	0.081	0.111	0.75	0.46

Table 6 Summary of the fixed effects in the second mixed-effects model for Experiments 1 and 2, data separated by CASE

	Accusative				Dative			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	-1.364	0.153	-8.95	2.93e-14***	-1.329	0.145	-9.14	1.11e-14***
ORDER	1.524	0.164	9.32	4.66e-15***	1.385	0.150	9.25	6.66e-15***

Table 7 Summary of the fixed effects in the third mixed-effects model for Experiments 1 and 2, data separated by ORDER

	SO				OS			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	0.160	0.042	3.86	2.06e-5***	-1.365	0.152	-8.96	2.71e-14***
CASE	-0.104	0.029	-3.59	5.32e-5***	0.037	0.049	0.76	0.449

Statistical results, self-paced reading time studies

Auxiliary position

See Table 8.

Table 8 Summary of the fixed effects in the mixed-effects model for Experiments 3 and 4 (self-paced reading times), auxiliary position

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.99	0.041	145.497	<2e-16***
CASE	0.007	0.016	0.47	0.636
ORDER	-0.026	0.015	-1.78	0.075
EXPERIMENT	-0.113	0.059	-1.91	0.059
CASE:ORDER	-0.012	0.021	0.59	0.558
CASE:EXP.	-0.0193	0.023	-0.85	0.395
ORDER:EXP.	-0.015	0.021	-0.69	0.490
CASE:ORDER:EXP.	0.034	0.030	1.13	0.260

First word post-auxiliary

See Tables 9, 10, 11, 12 and 13.

Table 9 Summary of the fixed effects in the mixed-effects model for Experiments 3 and 4 (self-paced reading times), first word following the auxiliary

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.912	0.025	235.62	<2e-16***
CASE	0.002	0.013	0.14	0.887
ORDER	-0.054	0.012	-4.69	2.79e-06***
EXPERIMENT	-0.085	0.036	-2.36	0.020*
CASE:ORDER	-0.009	0.016	-0.53	0.595
CASE:EXP.	-0.035	0.019	-1.86	0.064
ORDER:EXP.	-0.015	0.017	-0.92	0.356
CASE:ORDER:EXP.	0.045	0.023	1.92	0.055

Table 10 Summary of the fixed effects in the second mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), first word following the auxiliary

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.826	0.028	210.31	<2e-16***
CASE	-0.033	0.013	-2.62	0.0097**
ORDER	-0.069	0.014	-6.11	1.18e-09***
CASE:ORDER	0.037	0.016	2.28	0.023*

Table 11 Summary of the fixed effects in the mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), first word following the auxiliary

	Accusative				dative			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	5.826	0.028	209.68	<2e-16***	5.794	0.025	235.96	<2e-16***
ORDER	-0.069	0.012	-5.85	6.47e-09***	-0.033	0.011	-3.01	0.0027**

Data are separated by CASE

Table 12 Summary of the fixed effects in the mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), first word following the auxiliary

	SO				OS			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	5.758	0.024	239.06	<2e-16***	5.826	0.029	202.56	<2e-16***
ORDER	0.003	0.010	0.27	0.788	-0.033	0.012	-2.75	0.00599**

Data are separated by ORDER

Table 13 Summary of the fixed effects in the second mixed-effects model for Experiment 4 (self-paced reading times, COVERT stimulus set), first word following the auxiliary

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.912	0.024	251.17	<2e-16***
CASE	0.002	0.014	0.12	0.907
ORDER	-0.054	0.012	-4.49	7.62e-06***
CASE:ORDER	-0.009	0.017	-0.51	0.612

Second word post-auxiliary

See Table 14.

Table 14 Summary of the fixed effects in the mixed-effects model for Experiments 3 and 4 (self-paced reading times, OVERT and COVERT stimulus sets), second word following the auxiliary

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.874	0.026	230.617	<2e-16***
CASE	-0.010	0.012	-0.80	0.423
ORDER	-0.044	0.011	-4.05	5.18e-05***
EXPERIMENT	-0.089	0.037	-2.44	0.016*
CASE:ORDER	0.011	0.015	0.74	0.458
CASE:EXP.	0.002	0.017	0.12	0.903
ORDER:EXP.	0.014	0.016	0.92	0.357
CASE:ORDER:EXP.	-0.0001	0.022	-0.01	0.995

Third word post-auxiliary

See Table 15.

Table 15 Summary of the fixed effects in the mixed-effects model for Experiments 3 and 4 (self-paced reading times, OVERT and COVERT stimulus sets), third word following the auxiliary

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.887	0.025	237.97	<2e-16***
CASE	-0.008	0.010	-0.78	0.435
ORDER	-0.009	0.010	-0.96	0.335
EXPERIMENT	-0.071	0.036	-1.98	0.0498*
CASE:ORDER	0.005	0.014	0.37	0.715
CASE:EXP.	-0.017	0.015	-1.20	0.231
ORDER:EXP.	-0.001	0.014	-0.09	0.927
CASE:ORDER:EXP.	0.023	0.020	1.14	0.255

Fourth word post-auxiliary

See Tables 16, 17, 18, 19 and 20.

Table 16 Summary of the fixed effects in the mixed-effects model for Experiments 3 and 4 (self-paced reading times, OVERT and COVERT stimulus sets), sentence-final word

	Estimate	SE	t value	Pr(> t)
(Intercept)	6.003	0.033	179.58	<2e-16***
CASE	0.007	0.015	0.43	0.671
ORDER	0.013	0.014	0.93	0.355
EXPERIMENT	-0.027	0.048	-0.55	0.583
CASE:ORDER	-0.029	0.020	-1.49	0.136
CASE:EXP.	-0.033	0.022	-1.49	0.136
ORDER:EXP.	-0.045	0.020	-2.24	0.025*
CASE:ORDER:EXP.	0.073	0.028	2.59	0.0097**

Table 17 Summary of the fixed effects in the second mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), sentence-final word

	Estimate	SE	t value	Pr(> t)
(Intercept)	5.976	0.038	159.08	<2e-16***
CASE	-0.027	0.015	-1.83	0.067
ORDER	-0.032	0.014	-2.30	0.022*
CASE:ORDER	0.045	0.020	2.24	0.025*

Table 18 Summary of the fixed effects in the mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), sentence-final word

	Accusative				Dative			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	5.976	0.037	159.47	<2e-16***	5.949	0.034	173.62	<2e-16***
ORDER	-0.031	0.014	-2.20	0.028*	0.0120	0.014	0.86	0.392

Data are separated by CASE

Table 19 Summary of the fixed effects in the mixed-effects model for Experiment 3 (self-paced reading times, OVERT stimulus set), sentence-final word

	SO				OS			
	Estimate	SE	t value	Pr(> t)	Estimate	SE	t value	Pr(> t)
(Intercept)	5.944	0.036	165.66	<2e-16***	5.976	0.036	164.20	<2e-16***
CASE	-0.177	0.014	1.28	0.199	-0.025	0.014	-1.79	-0.073.

Data are separated by ORDER

Table 20 Summary of the fixed effects in the second mixed-effects model for Experiment 4 (self-paced reading times, COVERT stimulus set), sentence-final word

	Estimate	SE	t value	Pr(> t)
(Intercept)	6.003	0.031	195.67	<2e-16***
CASE	0.007	0.017	0.39	0.699
ORDER	0.013	0.0143	0.89	0.371
CASE:ORDER	-0.029	0.020	-1.45	0.147

Appendix B: Syntactic tests for the status of dative Case

In Appendix B, we give a list of tests that distinguish dative from accusative Case across various syntactic contexts. Most of these tests but not all of them have been discussed in Bayer et al. (2001) and certain previous publications. We indicate in each case how the result can be derived if datives involve the projection of a K(ASE)-phrase.

Function changing operations

Function changing operations suppress the subject's theta role and affect the verb's ability to license an object. If *v* is lacking, it is the accusative that is affected, not the dative. The dative is taken care of by the functional head K.

- (B1) a. Oswald hat den Präsidenten ermordet.
Oswald has the president.ACC assassinated
 'Oswald has assassinated the president.'
- b. Der Präsident wurde ermordet.
The president.NOM was assassinated
 'The president was assassinated.'
- (B2) a. Oswald hat dem Präsidenten gehuldigt.
Oswald has the president.DAT given.homage
 'Oswald gave homage to the president.'
- b. Dem Präsidenten/*der Präsident wurde gehuldigt.
the president.DAT was given.homage
 'The president was given homage.'

- (B3) a. Es ist leicht, diesen Wagen zu fahren.
It is easy this car.ACC to drive
 ‘It is easy to drive this car.’
- b. Dieser Wagen fährt sich leicht.
this car.NOM drives REFL easily
 ‘This car drives easily.’
- c. *Diesen Wagen fährt es sich leicht.
this car.ACC drives it REFL easily
- (B4) a. Es ist leicht, diesem Weg zu folgen.
it is easy this path.DAT to follow
 ‘It is easy to follow this path’.
- b. *Dieser Weg folgt sich leicht.
this path.NOM follows REFL easily
- c. Diesem Weg folgt es sich leicht.
this path.DAT follows it REFL easily
 ‘This path is easy to follow.’

Binding

Binding shows an asymmetry that disfavors dative arguments as potential binders. If the binding DP is in a KP-shell, it fails to c-command the anaphor in the same way as a DP inside a PP fails to c-command it.³¹

- (B5) a. Der Arzt₁ hat den Patienten₂ sich_{1/2} im Spiegel gezeigt.
the doctor has the patient.ACC REF in.the mirror shown
 ‘The doctor showed the patient himself in the mirror.’
- b. Der Arzt₁ hat dem Patienten₂ sich_{1/*2} im Spiegel gezeigt.
the doctor has the patient.DAT REF in.the mirror shown
 ‘The doctor showed himself to the patient in the mirror.’

In (B5a), either one of the arguments with structural Case may be the binder of the reflexive, whence the ambiguity. In (B5b), however, only the subject can bind the reflexive. The dative argument is in a KP-shell and does not c-command the reflexive, whence the non-ambiguity.

³¹ Importantly, this restriction pertains only to anaphoric (A-) binding, not to variable (A’-) binding.

- (i) Die Schwester hat jedem/ keinem Patienten seine Tabletten gegeben.
the nurse has each/ no patient-DAT his pills-ACC given
 ‘The nurse gave each/no patient his pills.’

The reason for this is that quantifier and negation features, etc., are part of the D- and therefore also of the K-system. See for discussion Bayer and Bader (2007) and references therein.

Secondary predication

As Vogel and Steinbach (1995) observe, the subject of a secondary predicate can be a nominative or an accusative but not a dative. This follows under the standard assumption that the subject is the external argument of the predicate and as such has to c-command it.

- (B6) a. Hans_x hat den Rektor_y schon dreimal betrunken_{x/y}
Hans.NOM has the rector.ACC already three.times drunk
 getroffen.
met
 ‘Hans met the vice chancellor drunk already three times.’ (ambiguous)
- b. Hans_x ist dem Rektor_y schon dreimal betrunken_{x/*y}
Hans.NOM is the rector.DAT already three.times drunk
 begegnet
encountered
 ‘Hans met the vice chancellor already three times while he, Hans, was drunk.’ (unambiguous)
- (B7) Hans_x hat mit dem Rektor_y schon dreimal betrunken_{x/*y}
Hans.NOM has with the rector.DAT already three.times drunk
 telefoniert.
telephoned
 ‘Hans telephoned with the vice-chancellor drunk already three times.’

While in (B6a), the predicate drunk can be ascribed either to Hans or to the vice chancellor, in (B6b) and in (B7), it can only be ascribed to Hans. The reason seems to be that in (B6b) the DP is embedded in a KP with K preventing c-command, and in (B7) it is embedded in a PP with P preventing c-command. As Williams (1994) showed, the latter contrast can also be found in English where (B8b) could at best mean that John was raw in the meat-eating event.

- (B8) a. John ate the meat_x raw_x
 b. *John ate at the meat_x raw_x

Extraction

Müller (1995) observes that dative as opposed to accusative DPs do not allow extraction:

- (B9) a. [Über Scrambling]₁ habe ich einem Buch über
about scrambling have I a book.DAT about
 Optionalität [einen Aufsatz t₁] hinzugefügt.
optionality an article.ACC added
 ‘I have added to a book about optionality an article about scrambling.’
- b. *[Über Optionalität]₂ habe ich einen Aufsatz über
about optionality have I an article.ACC about
 Scrambling [einem Buch t₂] hinzugefügt.
scrambling a book.DAT added
 ‘I have added an article about scrambling to a book about optionality.’

If K induces an extra functional layer for dative objects that is absent in accusative objects, the barrier effect is expected. Since in German PPs are an extraction island, the close analogy between datives and PP is highly suggestive:

- (B10) *[Über Scrambling]₁ habe ich stundenlang [in einem Buch t₁]
about scrambling have I hours.long in a book.DAT
 herumgeblättert.
turned.over.leaves
 ‘I have for hours browsed through a book about scrambling.’

Synthetic compounds

In synthetic compounds the object of the verb integrates (or incorporates) into the verb, but this integration is confined to direct objects which would be assigned accusative Case in syntax.

- (B11) a. Die Studentin betreut die Kinder regelmäßig.
the student looks.after the children.ACC regularly
 ‘The student takes care of the children regularly.’
- b. Mit Kinderbetreuen verdient man wenig.
With child.care earns one little
 ‘With child care one earns little.’
- (B12) a. Das Rote Kreuz hilft vielen Kindern.
the red cross helps many children.DAT
 ‘The Red Cross helps many children.’
- b. *Mit Kindernhelfen erlangt man selten Ruhm.
with child.help attains one rarely honor
 ‘Child care doesn’t gain you reputation.’

Incorporation being restricted to the X^o-domain, the object to be integrated must be non-phrasal. If datives are only licensed via a KP-shell, there is a reason why dative arguments resist integration in synthetic compounds.³²

Recoverability

Null arguments as they can be observed in topic-drop (alias ‘pronoun zap’) are recoverable by virtue of the functional structure that licenses Case. Since in German the unmarked projection of a transitive finite verb employs the functional heads *v* and T but nothing like AgrIO etc., we expect null arguments which recover nominatives and accusatives but no null arguments which would have to recover datives.

- (B13) a. Ich hab’ ihn schon gesehen.
I.NOM have him.ACC already seen
 ‘I saw him already.’
- b. [Ø]₁ hab’ t₁ ihn schon gesehen. [Ø]=NOM
- c. [Ø]₂ hab’ ich t₂ schon gesehen. [Ø]=ACC

PPs can never be dropped, even if their content is predictable from the verb, as could be the case in *denken (an)*, *nachdenken (über)*, *sich freuen (auf)*. Topic drop appears to be under tight control of the functional structure associated with the verb. Interestingly, the same is true for dative arguments. Consider the verbs *widersprechen* and *vertrauen*, both of which require a dative object.

- (B14) a. *[Ø]₂ widerspricht ja keiner t₂. [Ø]=DAT
objects PRT nobody
 ‘To him/her, nobody objects.’
- b. *[Ø]₂ würde ich t₂ nicht vertrauen. [Ø]=DAT
would I not trust
 ‘Him/her, I wouldn’t trust.’

The KP-hypothesis explains this straightforwardly by the fact that there is no K that could formally recover the empty topic.

Comparative clauses

Comparative clauses contain gaps, but they do so only when these gaps correspond to the structural Cases nominative or accusative.

³² We admit that it is unclear what the projective status of Case-less pronominals such as the reflexive *sich* should be. These are fine as objects of dative-assigning verbs as in *Er widerspricht sich* ‘He contradicts himself’. One possibility could be that these verbs can all function like inherently reflexive verbs, and that no Case is assigned at all. See however note 22.

- (B15) a. Mehr Patienten sind gekommen als [NOM ___] behandelt
more patients have come than treated
 werden konnten.
become could
 ‘More patients showed up than could be treated.’
- b. Mehr Patienten sind gekommen als der Arzt [ACC ___]
more patients are come than the doctor
 behandeln konnte.
treat could
 ‘More patients showed up than the doctor could treat.’
- c. *Mehr Patienten sind gekommen als der Arzt [DAT ___]
more patients are come than the doctor
 Medikamente verabreichen konnte.
medicine administer could
 ‘More patients showed up than the doctor could give medicine to.’

Interestingly, (B15c) can be rescued by insertion of the dative pronoun. The result is best when the object pronoun is scrambled into the Wackernagel position right after the finite verb.

- (B16) Mehr Patienten sind gekommen als [DAT ihnen]₁ der Arzt t₁ Medikamente verabreichen konnte.

These data show once again that datives rely on functional structure that cannot be derived from the verb but must be imported by the nominal expression itself.

View beyond

Although this list of tests may be impressive, we are far away from generalizations that go much beyond the German Case system. Cross-linguistic research in languages with a rich enough Case system is necessary to determine to what extent the role of dative, genitive and other oblique Cases bears resemblance to these findings. The fact that in Icelandic Case attraction appears to be prominent in datives but not in accusatives (see Wood et al. 2017) points in the direction of relatedness. Similarly, Kallulli (2016) concludes about Albanian that dative bare plurals are only seemingly bare, in the sense that there is morphological determiner drop which is however structurally present. Datives and locative pro-forms are always DPs in the sense that they contain a D-projection; this is not necessarily the case for direct objects. Bare singulars can only occur as direct objects, not as indirect objects.

Obviously, there are two issues, Case attraction and the morpho-syntactic representation of Case, that need to be brought together if we want to move toward a deeper exploration of the topic at hand.

References

- Alexiadou, Artemis, Paul Law, André Meinunger, and Chris Wilder. 2000. Introduction. In *The syntax of relative clauses*, ed. Artemis Alexiadou, Paul Law, André Meinunger, and Chris Wilder, 1–51. Amsterdam/Philadelphia: John Benjamins (= *Linguistik Aktuell/Linguistics Today* 32).
- Bader, Markus. 1996. *Sprachverstehen*. Opladen: Westdeutscher.
- Bader, Markus. 2012. The German *bekommen* passive: A case study on frequency and grammaticality. *Linguistische Berichte* 23: 249–281.
- Bader, Markus, and Josef Bayer. 2006. *Case and linking in language comprehension—Evidence from German*. Dordrecht: Springer (= *Studies in Theoretical Psycholinguistics* 34).
- Bader, Markus, Josef Bayer, Jens-Max Hopf, and Michael Meng. 1996. Case-Assignment in processing German verb-final clauses. In *Proceeding of the NELS 26 Sentence Processing Workshop*, 1–25 (= *MIT Occasional Papers in Linguistics* 9).
- Bader, Markus, and Jana Häussler. 2013. How much *bekommen* is there in the German *bekommen* passive? In *Non-canonical passives*, ed. Artemis Alexiadou and Florian Schäfer, 115–139. Amsterdam/Philadelphia: John Benjamins (= *Linguistik Aktuell/Linguistics Today* 205).
- Bader, Markus, and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28: 121–143.
- Bader, Markus, Michael Meng, and Josef Bayer. 2000a. Case and reanalysis 1. *Journal of Psycholinguistic Research* 29: 37–53.
- Bader, Markus, Michael Meng, Josef Bayer, and Jens-Max Hopf. 2000b. Syntaktische Funktionsambiguitäten im Deutschen—Ein Überblick. *Zeitschrift für Sprachwissenschaft* 19: 34–102.
- Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–278.
- Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48. <https://www.jstatsoft.org/article/view/v067i01>.
- Bayer, Josef. 1984. Comp in Bavarian syntax. *The Linguistic Review* 3: 209–274.
- Bayer, Josef, and Markus Bader. 2007. On the syntax of prepositional phrases. In *Interface and interface conditions*, ed. Andreas Späth, 157–179. Berlin: De Gruyter.
- Bayer, Josef, Markus Bader, and Michael Meng. 2001. Morphological underspecification meets Oblique Case: Syntactic and processing effects in German. *Lingua* 111: 465–514.
- Bayer, Josef, and Martin Salzmann. 2013. That-trace effects and resumption—How improper movement can be repaired. In *Repairs. The added value of being wrong*, ed. Patrick Brandt and Eric Fuß, 275–334. Boston/Berlin: De Gruyter Mouton.
- Behaghel, Otto. 1923–1932. *Deutsche Syntax*. vol. 4. Heidelberg: Winter.
- Béjar, Susana, and Milan Āezáč. 2009. Cyclic agree. *Linguistic Inquiry* 40: 35–73.
- Bianchi, Valentina. 2002a. Headed relative clauses in Generative Grammar—Part 1. *Glott International* 6 (7): 197–204.
- Bianchi, Valentina. 2002b. Headed relative clauses in Generative Grammar—Part II. *Glott International* 6 (8): 235–247.
- Bianchi, Valentina, and Christiano Chesi. 2014. Subject islands, reconstruction, and the flow of the computation. *Linguistic Inquiry* 45 (4): 525–569.
- Bittner, Maria, and Ken Hale. 1996. The structural determination of Case and agreement. *Linguistic Inquiry* 27: 1–68.
- Bock, Kathryn, Kathleen M. Eberhard, and J. Cooper Cutting. 2004. Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language* 51: 251–278.
- Bock, Kathryn, Kathleen M. Eberhard, J. Cooper Cutting, Antje S. Meyer, and Herbert Schriefers. 2001. Some attractions of verb agreement. *Cognitive Psychology* 43: 83–128.
- Bornkessel, Ina, Brian McElree, Matthias Schlesewsky, and Angela D. Friederici. 2004. Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language* 51: 495–522.
- Caha, Pavel. 2009. *The nanosyntax of case*. Doctoral Dissertation. Tromsø.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Cinque, Guglielmo. 2015. Three phenomena discriminating between “raising” and “matching” relative clauses. *Semantics-Syntax Interface* 2 (1): 1–27.

- Czypionka, Anna. 2014. *The interplay of object animacy and verb class in representation building*. Doctoral Dissertation, Humboldt-Universität zu Berlin.
- Czypionka, Anna, Katharina Spalek, Isabell Wartenburger, and Manfred Krifka. 2017. On the interplay of object animacy and verb type during sentence comprehension in German: ERP evidence from the processing of transitive dative and accusative constructions. *Linguistics* 55 (6): 1383–1433.
- De Vries, Marc. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry* 37: 229–270.
- Eberhard, Kathleen M., Kathryn Bock, and J. Cooper Cutting. 2005. Making syntax of sense: Number agreement in sentence production. *Psychological Review* 112 (3): 531–559.
- Fayol, Michel, Pierre Largy, and Patrick Lemaire. 1994. Cognitive overload and orthographic errors: When cognitive overload enhances subject-verb agreement errors: A study in French written language. *The Quarterly Journal of Experimental Psychology* 47 (A): 437–464.
- Fox, John and Sanford Weisberg. 2011. *An R companion to applied regression*. Thousand Oaks: Sage. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5: 519–559.
- Georgi, Doreen, and Martin Salzmann. 2014. Case attraction and matching in resumption in relatives. Evidence for top-down derivation. In *Topics at InfL*, ed. Anke Assmann, Sebastian Bank, Doreen Georgi, Timo Klein, Philipp Weisser and Eva Zimmermann, 347–395. Universität Leipzig (= *Linguistische Arbeitsberichte* 92).
- Georgi, Doreen, and Martin Salzmann. 2017. The matching effect in resumption: A local analysis based on Case attraction and top-down derivation. *Natural Language & Linguistic Theory* 35 (1): 61–98.
- Grimm, Jacob. 1866. Über einige Fälle der Attraction. In *Abhandlungen zur Literatur und Grammatik*, ed. I.I.I. Kleinere Schriften, 312–348. Berlin: Dümmler.
- Groos, Anneke, and Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar in theory of markedness in Generative Grammar. In *Proceedings of the 1979 GLOW Conference*, 171–216.
- Haider, Hubert. 1984. Mona Lisa lächelt stumm—Über das sogenannte deutsche “Rezipientenpassiv”. *Linguistische Berichte* 89: 32–42.
- Haider, Hubert. 1985. The case of German. In *Studies in German grammar*, ed. Jindrich Toman, 65–101. Dordrecht: Foris.
- Harbert, Wayne. 2007. *The Germanic languages*. Cambridge: Cambridge University Press.
- Häussler, Jana. 2009. *The emergence of attraction errors during sentence comprehension*. Doctoral dissertation, University of Konstanz.
- Heck, Fabian. 2005. *Gegen Kopfanhebung in deutschen Relativsätzen*. Handout GGS. Universität Tübingen. <http://home.uni-leipzig.de/heck/papiere/ggs05.pdf>.
- Heister, Julian, Kai-Michael Würzner, Johannes Bubenzer, Emund Pohl, Thomas Hanneforth, Alexander Geyken, and Reinhold Kliegl. 2011. dlexDB—eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau* 32: 10–20.
- Helgander, John. 1971. *The relative clause in English and other Germanic languages. A historical and analytical survey*. Doctoral Dissertation. University of Gothenburg.
- Hemforth, Barbara, Lars Konieczny, and Gerhard Strube. 1993. Incremental syntax processing and parsing strategies. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society (CogSci)* 15, 539–545.
- Hopf, Jens-Max, Markus Bader, Michael Meng, and Josef Bayer. 2003. Is human sentence parsing serial or parallel?: Evidence from event-related brain potentials. *Cognitive Brain Research* 15 (2): 165–177.
- Hopf, Jens-Max, Josef Bayer, Markus Bader, and Michael Meng. 1998. Event-related brain potentials and case information in syntactic ambiguities. *Journal of Cognitive Neuroscience* 10: 264–280.
- Kallulli, Dalina. 2013. (Non-)canonical passives and reflexives. Deponents and their like. In *Non-canonical passives*, ed. Artemis Alexiadou and Florian Schäfer, 337–358. Amsterdam/Philadelphia: John Benjamins (= *Linguistik Aktuell/Linguistics Today* 205).
- Kallulli, Dalina. 2016. Clitic doubling as differential object marking. *Rivista di Grammatica Generativa: Research in Generative Grammar (RGG)* 38: 161–171.
- Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8 (1): 63–99.
- Lamontagne, Greg, and Lisa Travis. 1987. The syntax of adjacency. In *Proceedings of the Sixth West Coast Conference on Formal Linguistics (WCCFL 6)*, 173–186.

- Lenerz, Jürgen. 1984. *Syntaktischer Wandel und Grammatiktheorie. Eine Untersuchung an Beispielen aus der Sprachgeschichte des Deutschen*. Tübingen: Niemeyer.
- McGinnis, Martha. 1996. Projection and position. In *Proceedings of ConSole IV*, 203–220.
- Meng, Michael, and Markus Bader. 2000. Ungrammaticality detection and garden-path strength: Evidence for serial parsing. *Language and Cognitive Processes* 15: 615–666.
- Müller, Gereon. 1995. *A-bar syntax. A study of movement types*. Berlin: Mouton de Gruyter (= *Studies in Generative Grammar* 42).
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiiian, and Wendy K. Wilkins, 262–294. Amsterdam: John Benjamins. (= *Linguistik Aktuell/ Linguistics Today*, 101).
- Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34: 37–90.
- Pittner, Karin. 1995. The case of German relatives. *The Linguistic Review* 12: 197–231.
- Pittner, Karin. 1996. Attraktion, Tilgung und Verbposition. Zur diachronen und dialektalen Variation beim Relativpronomen im Deutschen. In *Language change and Generative Grammar*, ed. Ellen Brandner and Gisella Ferraresi, 120–153. Wiesbaden: Westdeutscher Verlag (= *Sonderheft Linguistische Berichte* 7).
- R Development Core Team. 2005. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Rohde, Doug. 2003. *Linger: A flexible platform for language processing experiments* (Version 2.94). <http://tedlab.mit.edu/Linger>.
- Salzmann, Martin. 2006. *Resumptive prolepsis. A study in indirect A'-dependencies*. Doctoral Dissertation, University of Leiden. (= *LOT Dissertation Series* 136).
- Schlesewsky, Matthias, Stefan Frisch, and Ina Bornkessel. 2001. Das Problem mit syntaktischen Funktionsambiguitäten: Eine kritische Betrachtung zu einem Überblick von Bader/Meng/Bayer/Hopf. *Zeitschrift für Sprachwissenschaft* 20: 251–265.
- Schriefers, Herbert, Angela D. Friederici, and Katja Kuhn. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language* 34 (4): 499–520.
- Tremblay, Antoine, and Johannes Ransijn. 2015. *Lmerconveniencefunctions: Model selection and post hoc analysis for (g)lmer models*. <https://CRAN.R-project.org/package=LMErConvenienceFunctions>. R (package version 2.10).
- van Riemsdijk, Henk. 1989. Swiss relatives. In *Sentential complementation and the lexicon*, ed. Dany Jaspers, Wim Klooster, Yvan Putseys, and Pieter Seuren, 343–354. Dordrecht: Foris.
- Vogel, Ralf, and Markus Steinbach. 1995. On the (absence of a) base position for dative objects in German. *FAS Papers in Linguistics* 4: 99–131.
- Wagers, Matthew W., Ellen F. Lau, and Colin Phillips. 2009. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* 61 (2): 206–237.
- Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12): 1–20. <http://www.jstatsoft.org/v21/i12/paper>.
- Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis*. New York: Springer. <http://ggplot2.org>.
- Wickham, Hadley. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* 40(1): 1–29. <http://www.jstatsoft.org/v40/i01/>.
- Williams, Edwin. 1994. *Thematic structure in syntax*. Cambridge, MA: MIT Press.
- Wood, Jim, Einar Freyr Sigurðsson, and Iris Edda Nowenstein. 2017. Inverse attraction in Icelandic relative clauses. In *Syntactic variation in Insular Scandinavian*, ed. Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen, and Zakaris Svabo Hansen, 199–232. Amsterdam/Philadelphia: John Benjamins. (= *Studies in Germanic Linguistics* 1).
- Woolford, Ellen. 2006. Lexical Case, inherent Case, and argument structure. *Linguistic Inquiry* 37 (1): 111–130.