

# Gradient Acceptability of Case Conflicts as Grammatical Illusion

Ralf Vogel\*

“... many factors influence relative acceptability and the relative frequency of a linguistic item. [...] The controversy [...] is whether it makes sense to keep at least some of these factors outside of what one specifies in one’s grammar, and if so, whether one can keep *all* factors that introduce gradient properties external to grammar. (Fanselow et al. 2006: 4f)

[...] the gradient nature of acceptability need not reflect any gradience of grammaticality. How can we decide if grammar is categorical or not? Of course, this question is answered by the identification of the best, i.e. most predictive grammatical model, but there are some hints at an answer. If a gradient difference in acceptability can be shown to disappear when [...], gradience has been eliminated from grammar. (Fanselow 2007: 365f)

Gisbert Fanselow, together with his team and colleagues at Potsdam University, played a central role in developing methods of empirical linguistics for grammatical theory. One of the problems that figured prominently in his contributions is the one that he identified in the quotations given above, namely, that elicited data are often gradient, despite the non-gradience of single grammatical rules: either the rules and constraints of grammar are there or they are not there, either they are respected or not. The strongest hypothesis

---

\*I want to thank the organisers of the 2023 GGS workshop on Gisbert’s contribution to syntactic theory, as well as the editors of this volume, especially Fabian Heck. I am also grateful to the audience of the workshop for a fruitful discussion and many helpful suggestions. I also benefitted from the discussion of these ideas at the Bielefeld linguistics colloquium in summer 2023.

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – CRC-1646 “Linguistic Creativity in Communication”, project number 512393437, project A04 “Empirical Profiles of grammatical creativity”.

*Gisbert Fanselow’s Contributions to Syntactic Theory*, 47–58

Artemis Alexiadou, Doreen Georgi, Fabian Heck, Gereon Müller & Florian Schäfer (eds.)

LINGUISTISCHE ARBEITS BERICHTS 96, Universität Leipzig 2024

that follows from this perspective, then, is that the gradience that is ubiquitous in empirical studies has no grammar-internal cause.

In Fanselow (2007), the external factors Gisbert Fanselow discussed are information structure and contextual fit, prosodic complications, and relative ease of processing, all of which indeed are syntax-external factors that could degrade the acceptability rating of linguistic structures without inviting the conclusion that gradience was an intrinsic feature of grammar.

From my own work, I can add prescription bias (Vogel 2019) and the incompleteness of grammatical rule systems (Vogel 2023, see also Reis 1979, 2017) as further factors. An important methodological aspect lies in the clarification of the question when gradience is grammatically relevant. My (Vogel 2019) more conservative approach is that quantitative contrasts have to be of a particular size ('medium' effect size in terms of Cohen 1988, 1992) in order to be taken as grammatically relevant. This leads to a significant reduction of the number of relevant gradience phenomena.

In the current paper, I will propose a further factor, *grammatical illusions* in the sense of Fanselow and Frisch (2006) and Haider (2011). The illusions I am talking about, though, are the illusions of both speakers and linguists.

Before going into details, one disclaimer is in order: my personal view on this issue is that *markedness* is a grammar-internal factor that, under particular circumstances, causes degraded acceptability. It does not follow from this, in my view, that grammar is gradient. But I am certain that the most relevant answers to the question whether there is gradience in grammar will emerge from studying the correlation of grammatical markedness and patterns of gradience in quantitative studies.

The structural vs. oblique case distinction, which is at the centre of my discussion here, is such a markedness distinction. Structural case is unmarked, oblique case is marked. In German, my main subject language, the structural cases are nominative, mainly the case of the grammatical subject, and accusative, the default case of grammatical objects. The main oblique case in the sentential domain is dative.

The distinction has syntactic consequences without yet introducing gradient grammaticality. The text book from which I learned linguistics as a student was Fanselow and Felix (1987), which was mainly written by Gisbert Fanselow. On pages 84ff of that book, a very clear account of the structural/oblique case distinction for German is given, based on Stowell (1981). The following assumptions are being made:

- only NPs can have case
- sentences cannot have case
- structural case can remain unrealised
- oblique case cannot remain unrealised

One empirically correct consequence of these assumptions is that subordinate clauses can occur in place of NPs with structural, but not with oblique case. Of course, this is about English and German:<sup>1</sup>

- (1) a. Hans leugnete ...  
 H. denied  
 (i) daß er das Auto gestohlen hat  
 that he the car stolen has  
 (ii) den Diebstahl (des Autos)  
 the.ACC theft of the car  
 (iii) das Auto gestohlen zu haben  
 the car stolen to have
- b. Hans weiß ...  
 H. knows  
 (i) welche Telefonnummer Maria hat  
 which phone number M. has  
 (ii) Marias/die Telefonnummer  
 M.'s/the.ACC phone number  
 (Fanselow and Felix 1987: 85)

(1) illustrates the situation for accusative case. (2) shows the parallel facts for nominative case:

- (2) a. Der Diebstahl wurde von Hans geleugnet  
 the.NOM theft was by H. denied  
 b. Dass er das Auto gestohlen hat, wurde von Hans geleugnet  
 that he the car stolen has was by H. denied  
 c. Das Auto gestohlen zu haben, wurde von Hans geleugnet  
 the car stolen to have was by H. denied

So nominative and accusative are the two structural cases of German and they

---

<sup>1</sup>Note that both *leugnen* and *wissen* assign accusative to an object noun phrase.

share the syntactic properties that result from this status. The situation for dative case is different:

- (3) a. Die Darstellung entspricht nicht den Tatsachen  
 the presentation conforms not the.DAT facts  
 b. \*Die Darstellung entspricht nicht, daß dieser Verlust uns so  
 the presentation conforms not that this loss us so  
 schwer traf  
 hard hit  
 (Fanselow and Felix 1987: 85)

The claim that clausal complements are impossible in place of the oblique dative case is corroborated. The NP arguments in these examples denote propositional content, just like the clausal complements do, and are therefore equivalent with respect to semantic roles.

Case-related phenomena from German that do introduce gradient acceptability are *case conflicts in free relative clauses (FR)*. Their acceptability patterns suggest a markedness scale like *nom* < *acc* < *dat*. And this could translate into gradient grammaticality.

It appears to be the case that the *wh*-pronoun in the initial position of a free relative clause (not only in German) is subject to two independent, and therefore potentially conflicting case requirements, that of the verb in the relative clause, and that of the matrix verb. Consider the examples in (4).

- (4) a. Ich besuche, wen ich mag  
 I visit→ACC who-ACC I like→ACC  
 b. Ich besuche wem ich vertraue  
 I visit→ACC who-DAT I trust→DAT  
 c. Ich besuche wer mir vertraut  
 I visit→ACC who-NOM I-DAT trust→NOM  
 d. Ich besuche wen ich vertraue  
 I visit→ACC who-ACC I trust→DAT

Linguists have disagreed in their assessments of these patterns. Groos and van Riemsdijk (1981) claimed that only (4-a) was grammatical. Their claim is that the two cases have to match in German. Pittner (1991) argued that (4-a,b) were grammatical. Her explanation was that (4-b) is in compliance with the *case hierarchy*: *nom* < *acc* < *dat* such that the case that is ignored (accusative)

is ‘lower’ in the hierarchy (less marked) than the case that surfaces on the *wh*-pronoun (dative).

An additional factor is, of course, that in German the FR pronoun must occur with the case that is required within the relative clause, and this factor is mainly responsible for the ungrammaticality of (4-d).

I claimed in Vogel (2001) that (4-a,b,c) were grammatical. Vogel (2011) provided empirical evidence in this respect and showed that acceptability decreases from (4-a) to (4-d). I further suggested that perhaps all scholars may agree on this.

I tested the patterns in (4) again in a recent acceptability rating study. In this study, subjects were presented test sentences as auditory stimuli (twice in a sequence) before they were prompted to provide their acceptability rating on a four-point rating scale (✓ = ‘unproblematic’; ? = ‘grammatical but slightly problematic’; ?? = ‘problematic, but not ungrammatical’; \* = ‘ungrammatical’), within a time window of eight seconds. The stimuli were presented with 8 different voices (4 male, 4 female). Each of the four conditions of the sub-experiment on FRs was presented in eight different lexical variants, whereby each subject heard four different items per condition. Subjects heard 190 sentences overall, 130 test sentences from 13 different experiments and 60 fillers. 94 persons (students from Bielefeld University, 47 male/female) took part in this experiment. Table 1 displays the distribution of subjects’ median ratings.<sup>2</sup>

		*	??	?	✓	median			
≈(4-a)	acc ACC	1	0	6	4	22	9	52	✓
≈(4-b)	acc DAT	7	6	21	10	16	9	25	?
≈(4-c)	acc NOM	21	6	31	8	15	6	7	??
≈(4-d)	ACC dat	47	10	17	8	11	0	1	*

Table 1: Distribution of subjects’ median ratings for four case conflicts in FRs

<sup>2</sup>Because of the even number of judgements per condition, subjects median ratings could fall between two rating levels, therefore there are seven different rating levels in Table 1. The median subject for (4-d), strictly speaking, is between the lowest level ‘\*’ and the second lowest level that lies between ‘\*’ and ‘??’. In the second column in Table 1, the case patterns are given with the matrix case preceding the FR case and capital letters indicating the case form of the FR pronoun.

The distribution of subjects' median ratings pretty much confirms my claim from Vogel (2011). More specifically, we can ask whether individual speakers also make these distinctions. Table 2 displays how subjects are distributed when we look at their relative assessments of the four conditions in (4), comparing each pattern with the one at the next lower acceptability level. The first row, for instance, states that 61 subjects rated (4-a) better than (4-b), in line with the linguists, 30 gave them equal rating, and 3 had a higher rating for (4-b).

	>	=	<
≈(4-a) vs. (4-b)	61 (65%)	30 (32%)	3 (3%)
≈(4-b) vs. (4-c)	57 (61%)	25 (27%)	12 (13%)
≈(4-c) vs. (4-d)	51 (54%)	27 (29%)	16 (17%)

Table 2: Distribution of subjects relative assessments of the three central comparisons of case conflicts in FRs

For each of the three cases, Table 2 displays absolute majorities in the predicted direction (the '>' column), for the contrast between (4-c) and (4-d) it is less robust than in the other two cases. In order to test these distributions for effect size, I calculated the measure of *Cliff's delta* (Cliff 1996)<sup>3</sup> for meeting the threshold of a medium effect size, which has been proposed to lie at the level of 0.333 (Romano et al. 2006). Cliff's delta is above 0.333 for each of the three contrasts, and only the confidence interval for the contrast between (4-c) and (4-d) ranges clearly below that threshold, see Table 3.

The experimental evidence for this four-level distinction of gradient acceptability is thus fairly robust. Given what I discussed above about accusative case, this comes as a surprise. Consider again that the FRs tested here stand in for accusative objects:

- (5) Ich besuche [ **accusative object** ]  
I visit

<sup>3</sup>Cliff's delta for comparing two experiment conditions X and Y is here calculated by dividing, for each row in Table 2, the difference between counts in the second and the fourth cell by the row sum (= 94 here). E.g., for (4-a) vs (4-b):  $\text{delta} = (61-3)/94 = 0.617$ . The confidence intervals are calculated with the statistics software *R* (R Core Team 2016) and the R package *orddom* (Rogmann 2013).

Cliff's delta	
≈(4-a) vs. (4-b)	0.617 [0.492;0.717]
≈(4-b) vs. (4-c)	0.479 [0.321;0.611]
≈(4-c) vs. (4-d)	0.372 [0.201;0.516]

Table 3: Cliff's delta with 95% confidence intervals for subject's medians for the three relevant comparisons of case conflicts in FRs

Recall from Fanselow and Felix (1987) that accusative is structural case, and that accordingly subordinate clauses should be possible as accusative objects anyway. FRs look like ordinary *wh*-clauses. These can also be used in place of accusative objects, as shown in (1-b-i). So where does the gradient acceptability in (6-ab), repeated from (4), come from?<sup>4</sup>

- (6) a. ??Ich besuche wer mir gefällt  
 I visit→ACC who.NOM me.DAT pleases  
 b. ?Ich besuche wem ich vertraue  
 I visit→ACC who.DAT I trust

The matching effect in FR constructions is attested for many languages. It is often treated with ad hoc analyses of a hidden [NP – CP] structure of FRs. This leads to a special syntactic analysis of sentences that often look like ordinary *wh*- or relative clauses. The ad hoc nature of these analyses is obvious and I therefore resisted to take such a route. For German, it has always been clear that the matching effect is rather superficial. The inanimate *wh*-pronoun *was* is homophonous for nominative and accusative. We showed in Vogel et al. (2006) that the degraded acceptability rating for a case conflict like (6-a) disappears in cases with inanimate *was*, as illustrated in (7), where (7-a) is matching and (7-b) is non-matching.

- (7) a. Ich esse was ich mag  
 I eat→ACC what.ACC I like→ACC  
 b. Ich esse was mir gefällt  
 I eat→ACC what.NOM me pleases→NOM

<sup>4</sup>The reported judgements are the median ratings from the experiment.

The standard explanation for this contrast between the ratings for (7) as compared to those for (6) is based on the assumption of a kind of ‘non-gradient acceptability illusion’ with respect to (7-b), such that the syncretism of *was* hides the underlying case conflict. The ratings in (6) are taken as the normal case. But it could be exactly the other way around: it is the non-matching clause-initial *wh*-pronouns that produce an irritation during parsing, a kind of inverse garden path effect with a temporary, local incoherence that occurs within an overall unproblematic syntactic configuration. That irritation is missing in (7-b), and as that sentence is unproblematic in principle, it is rated as fully acceptable.

According to that reasoning, the matching effect is a grammatical illusion (Fanselow and Frisch 2006, Haider 2011), but one of a special kind: the illusion of the degraded acceptability (not ungrammaticality) of non-matching FRs in place of structural case. Experiment participants as well as linguists react in their ratings on a superficially salient, but in fact irrelevant property of the FR, the case marking of the initial *wh*-pronoun. The gradient acceptability is then explained as a processing effect and the standard account about structural case in German from Fanselow and Felix (1987) prevails.<sup>5</sup>

But what about oblique case? Remember that oblique case cannot be realised by a subordinate clause. Therefore, FR clauses, not bearing case according to what I just concluded, should not be possible here. (8-b) should be as unacceptable as (8-a), repeated from (3-a), and (8-d). Surprisingly, it patterns with (8-c).

- (8) a. \*Die Darstellung entspricht nicht, daß dieser Verlust uns so  
       the presentation conforms not that this loss us so  
       schwer traf  
       hard hit
- b. Ich helfe wem ich vertraue  
       I help→DAT who.DAT I trust→DAT

---

<sup>5</sup>I am grateful to Fabian Heck (p.c.) for hinting me at an analysis by Asarina (2013) on case nominative-accusative syncretisms in Russian which makes a similar point as suggested here: syncretic forms for these two structural cases are “neutral” in the sense that their featural make-up is fully compatible with both cases simultaneously, and therefore resolve case conflicts. Translated to German, this would be the case for subordinate clauses in structural case positions. Such an account does not resolve the issue of oblique dative case discussed below.



- c. Ich helfe dem Freund  
 I help→DAT the.DAT friend
- d. \*Ich helfe wen ich mag  
 I help→DAT who.ACC I like→ACC

This is the sole problem that remains. Note also that the alternative would have similar problems: assume that contrary to what we just concluded the *wh*-pronoun of an FR serves as the case exponent of the FR and the FR thus has case. That would imply that (9-a) is structurally equivalent to (9-b) and should also be rated as ungrammatical, contrary to fact.

- (9) a. ?Ich besuche wem ich vertraue  
 I visit→ACC who.DAT I trust
- b. \*Ich besuche dem Freund  
 I visit→ACC the.DAT friend

So, no matter how we phrase our analysis, we seem to enter into contradictions.

On the other hand, there is no general ban on wellformedness illusions to become conventionalised, so that (8-b) may have become grammatical in the language community only for the reason that it is an illusion that *everyone* once had. For that reason, it is no illusion anymore, but an ordinary rule of the (E-)language. Still, there is no gradience – sentences like (8-b) receive quite high ratings.

Does this solve the problem? Apparently yes. There is no more gradience to be explained in grammatical terms. But this is achieved only by damaging main underlying assumptions. The idea of a non-gradient grammar is motivated by the idea of a “clean” or “pure” underlying grammar. If illusions can become rules of the grammar, this no longer holds. Yet, a grammar full of illusions could still be non-gradient.

German is one particular case in the quite diverse typology of this phenomenon. Languages that display case attraction go beyond the German pattern in that the FR pronoun surfaces with the case required by the matrix verb, even in non-matching configurations. Is an illusion account feasible even for such languages? This question needs to be addressed by taking into account the general regulations for case realisation in the respective languages. Case attraction may then either instantiate an illusion as well, or these patterns indeed have a different syntax of FRs or at least different regulations for the

morphology-syntax interface. Addressing these issues of course lies beyond the limits of the current paper.

The exercise that I carried out in this paper has another more general message which, I think, is pretty much in the spirit of how Gisbert Fanselow thought about the relation between grammatical theory and experimental linguistics. From his writings on these issues, you can get the very clear message that he neither agreed with the standard allegation by empirical linguists about the ‘armchair’ linguist’s introspective ‘(non-)evidence’, nor did he agree that linguists had some kind of privileged access to linguistic facts. More often than not, linguists and non-linguists (whom I would refuse to sort into experts and lay persons – they are all experts, as long as they speak the language) agree in their judgements.

What I want to suggest here, adding to these considerations, is that linguists and non-linguists may also agree in their illusions and erroneous views on the grammatical state of affairs of their common language, and this, of course, would be highly relevant for grammatical analysis, linguistic theory and methodology.

## References

- Asarina, Alya (2013): Neutrality vs. Ambiguity in Resolution by Syncretism: Experimental Evidence and Consequences. *In*: Y. Feinleib, N. LaClara and Y. Park, eds., *NELS 41: Proceedings of the Forty-First Annual Meeting of the North East Linguistic Society*. Vol. I, Graduate Linguistics Student Association, University of Massachusetts, Amherst, MA, pp. 43–56.
- Cliff, Norman (1996): *Ordinal methods for behavioral data analysis*. Routledge, London.
- Cohen, Jacob (1988): *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn, Erlbaum, Hillsdale, NJ.
- Cohen, Jacob (1992): ‘A Power Primer’, *Psychological Bulletin* **112**(1), 155–159.
- Fanselow, Gisbert (2007): ‘Carrots – perfect as vegetables but please not as main dish’, *Theoretical Linguistics* **33**, 353–367.
- Fanselow, Gisbert, Caroline Féry, Matthias Schlesewsky and Ralf Vogel (2006): *Gradience in Grammar: Generative Perspectives*. Oxford University Press, Oxford.
- Fanselow, Gisbert and Sascha Felix (1987): *Sprachtheorie. Band 2: Die Rektions- und Bindungstheorie*. UTB Francke, Tübingen.
- Fanselow, Gisbert and Stefan Frisch (2006): Effects of Processing Difficulty on Judgments of Acceptability. *In*: G. Fanselow, C. Féry, M. Schlesewsky and R. Vogel,

- eds., *Gradience in Grammar: Generative Perspectives*. Oxford University Press, Oxford, pp. 291–316.
- Groos, Anneke and Henk van Riemsdijk (1981): Matching Effects with Free Relatives: a Parameter of Core Grammar. In: A. Belletti, L. Brandi and L. Rizzi, eds., *Theories of Markedness in Generative Grammar*. Scuola Normale Superiore di Pisa, Pisa, pp. 171–216.
- Haider, Hubert (2011): ‘Grammatische Illusionen – Lokal wohlgeformt – global deviant’, *Zeitschrift für Sprachwissenschaft* **30**, 223–257.
- Pittner, Karin (1991): Freie Relativsätze und die Kasushierarchie. In: E. Feldbusch, ed., *Neue Fragen der Linguistik*. Niemeyer, Tübingen, pp. 341–347.
- R Core Team (2016): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.  
**URL:** <https://www.R-project.org/>
- Reis, Marga (1979): Ansätze zu einer realistischen Grammatik. In: K. Grubmüller, ed., *Befund und Bedeutung. Zum Verhältnis von Empirie und Interpretation in Sprach- und Literaturwissenschaft*. Niemeyer, Tübingen.
- Reis, Marga (2017): Grammatische Variation und realistische Grammatik. In: M. Konopka and A. Wöllstein, eds., *Grammatische Variation. Empirische Zugänge und theoretische Modellierung*. de Gruyter, Berlin and Boston, pp. 255–282.
- Rogmann, Jens J. (2013): *orddom: Ordinal Dominance Statistics*. University of Hamburg, department of psychology, Hamburg, Germany. R package version 3.1.  
**URL:** <https://CRAN.R-project.org/package=orddom>
- Romano, Jeanine, Jeffrey D. Kromrey, Jesse Coraggio, Jeff Skowronek and Linda Devine (2006): ‘Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen’s d indices the most appropriate choices?’, Annual meeting of the Florida Association of Institutional Research, Arlington, Virginia Oct 14-17, 2006.
- Stowell, Tim (1981): *Origins of Phrase Structure*. PhD thesis, MIT, Cambridge, Massachusetts.
- Vogel, Ralf (2001): Case Conflict in German Free Relative Constructions. An Optimality Theoretic Treatment. In: G. Müller and W. Sternefeld, eds., *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 341–375.
- Vogel, Ralf (2011): Disagreement, variation, markedness and other apparent exceptions. In: H. J. Simon and H. Wiese, eds., *Expecting the unexpected: exceptions in grammar*. de Gruyter Mouton, Berlin, pp. 339–359.
- Vogel, Ralf (2019): ‘Grammatical taboos. An investigation on the impact of prescription in acceptability judgement experiments’, *Zeitschrift für Sprachwissenschaft* **38**(1), 37–79.
- Vogel, Ralf (2023): Grammatical gaps, grammatical invention and grammatical theory. In: T. Strobel and H. Weiß, eds., *Grammatical gaps: definition, typology and theory*.

Vol. 34 of *Linguistische Berichte Sonderheft*, Helmut Buske Verlag, Hamburg, pp. 15–49.

Vogel, Ralf, Stefan Frisch and Marco Zugck (2006): ‘Case Matching. An Empirical Study On The Distinction Between Abstract Case And Case Morphology’, *Linguistische Berichte* **208**, 357–383.