

# Some qualms about orders and movements

Ralf Vogel

Bielefeld, the place where I have been living for 16 years now, is typical of the larger cities in the north-west of Germany. As an industrial centre, it was hit hard in the Second World War. The rebuilding of its infrastructure in the post-war period until the 1980s or so served the automotive industry, which had been pushed with especially high priority, it occurs to me, in the state of North Rhine-Westphalia, where car manufacturing has been a substantial part of the economy.

From the late nineteenth century until before WWII, Bielefeld was one of the centres of bicycle manufacturing in Germany.<sup>1</sup> But not even that could soften the city's car-centred rebuilding after the war.

Times have changed, however: the city council has rediscovered the bike. And so we find the construction of bike lanes all over the city.<sup>2</sup> But finding appropriate space for bike lanes under such conditions appears to be an art in itself. The typical bike lane in Bielefeld is small and switches constantly between sidewalk and street.

A biker in Bielefeld is always in the way of either other bikers, pedestrians or cars. So, as a biking syntactician in this city you are constantly in the middle of your profession: you never stop wondering about the rules that govern who is allowed to move first, where movement is blocked and why, what happens

---

*Strict Cycling: A Festschrift for Gereon Müller, 477–488*

Silke Fischer, Doreen Georgi, Fabian Heck, Johannes Hein, Anke Himmelreich, Andrew Murphy & Philipp Weisser (eds.)

STRICT CYCLING, Universität Leipzig 2024

<sup>1</sup>If I recall it correctly, I learned to ride a bike on a *Göricke* bike made in Bielefeld. Another aspect of Bielefeld's bicycle history is that cycling has been promoted early on in the region. Bielefeld's first cycling track was among the first ones in Germany, when it was opened in 1885. The Bielefeld *Radrennbahn*, still among the fastest cycling tracks in Europe, was opened in 1953 and is a protected monument today. Once or twice a year, you can still watch the spectacle of *Steher* races, where racing cyclists drive in high speed in the slipstream of a pacemaker motor bike – another product Bielefeld's industry was famous for until 1960.

<sup>2</sup>Bielefeld has a strong “pro bicycle”, i.e. pro bike lane, movement that has the support of the current city council. I suppose that bikers would be better off without bike lanes in Bielefeld, just riding on the roads, accompanied by a city-wide 30 km/h speed limit.

when movement paths intersect and so on. And you absolutely learn your lesson about the benefits of parallel movement (Müller 2001) on bike lanes.

To make things worse, the city council has installed barriers here and there on bike lanes, wherever they felt they needed to do bikers a favour. These barriers can only be crossed by one bike or person at a time. When two arrive simultaneously at such a barrier, as in Figure 1, the set of constraints governing who drives first appears to be as multi-faceted as the constraints governing the order in binomials in Müller (1997): I let pedestrians, children and old people go first, I won't let adult women on bikes go first (this no longer appears to be politically correct). If none of these criteria help, the one that was earlier (even only slightly) may go first, but if even that doesn't help, we get stuck and resolution may take ages, as both will be reluctant to go first (and appear impolite).

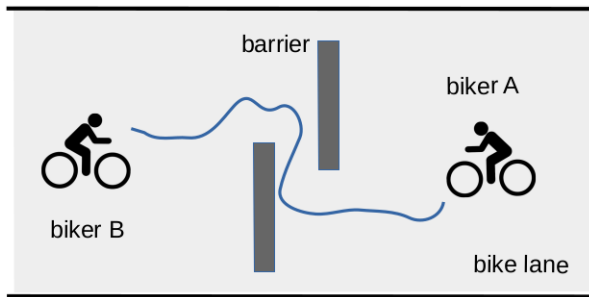


Figure 1: Barriers blocking movement on Bielefeld's bike lanes

The syntactician's moral: true optionality may not always be helpful. Constraints sometimes make life easier. Optionality can, of course, serve as a diagnostic tool: it is indicative of the *absence* of constraints.

But this may also be a matter of theoretical preferences. Pullum & Scholz (2001) (see also Pullum 2013) introduced the distinction between two ways of formulating theories of syntax and of explicating the grammar of a language, a *model-theoretic* and a *generative-enumerative* approach. Their differences can be illustrated with the problem sketched in Figure 1.

Suppose that there are constraints like *old people first*, *children first*, *first in – first out* etc., but that none of these apply here, as biker A and biker B are of the same age and arrived at the barrier simultaneously. So there is no constraint that would require either biker A or biker B to cross the barrier first.

The *model-theoretic* perspective would simply set the bikers free: whichever solution (‘biker A first’ or ‘biker B first’) you may choose, it will be fine, as neither solution would violate a constraint. The model-theoretic approach is thus somewhat agnostic about how the two bikers arrive at a solution. It is not only a non-derivational perspective, but it also abstracts away from details of execution. Table 1 may be an OT version of this, a really trivial account, I admit.

	biker A $\approx$ biker B	KIDS-1 <sup>ST</sup>	OLDIES-1 <sup>ST</sup>	1 <sup>ST</sup> IN-1 <sup>ST</sup> OUT
☞	biker A 1 <sup>st</sup>	✓	✓	✓
☞	biker B 1 <sup>st</sup>	✓	✓	✓

Table 1: ‘model-theoretic’ OT solution

Table 1 implies a kind of *underdetermination* of the grammar: not every aspect in which candidates may differ is subject to a constraint. Crucially, the input is not used to make a difference between similar candidates via faithfulness constraints either.

From the perspective of a *generative-enumerative* approach, as classified by Pullum & Scholz (2001) and Pullum (2013), the situation is quite different, because two candidates can only be non-identical if there has been some difference between them in their derivational history. Furthermore, as every derivational step in such a procedure must be defined within the system, it is impossible to remain agnostic about such differences in the same way as in a model-theoretic account. OT can of course be used within the derivation as well (see e.g. Heck & Müller 2013).

But it appears to me that a central aspect of a generative-enumerative OT variant should be that there may only be one winner per competition. This can be achieved, in principle, in two ways, via markedness or faithfulness.

The markedness approach would work with two tied ad hoc constraints favouring exactly one of the two options. The two constraints are assumed to be globally tied, i.e. we have in fact two parallel grammars with two different constraint rankings. These are given in Table 2. Each of these parallel grammars has now exactly one winner. The faithfulness variant would use the input to implement the contrast and then let faithfulness decide, see Table 3.

The only purpose of the two additional constraints in Table 2 is to make a

biker A $\approx$ biker B		KIDS-1 <sup>ST</sup>	OLDIES-1 <sup>ST</sup>	1 <sup>ST</sup> IN-1 <sup>ST</sup> OUT	A-1 <sup>ST</sup>	B-1 <sup>ST</sup>
☞	biker A 1 <sup>st</sup>	✓	✓	✓	✓	*
	biker B 1 <sup>st</sup>	✓	✓	✓	*	✓
biker A $\approx$ biker B		KIDS-1 <sup>ST</sup>	OLDIES-1 <sup>ST</sup>	1 <sup>ST</sup> IN-1 <sup>ST</sup> OUT	B-1 <sup>ST</sup>	A-1 <sup>ST</sup>
☞	biker A 1 <sup>st</sup>	✓	✓	✓	*	✓
	biker B 1 <sup>st</sup>	✓	✓	✓	✓	*

Table 2: ‘markedness’ 1-winner solution

biker A $\approx$ biker B		KIDS-1 <sup>ST</sup>	OLDIES-1 <sup>ST</sup>	1 <sup>ST</sup> IN-1 <sup>ST</sup> OUT	FAITH
A-1 <sup>st</sup>					
☞	biker A 1 <sup>st</sup>	✓	✓	✓	✓
	biker B 1 <sup>st</sup>	✓	✓	✓	*
biker A $\approx$ biker B		KIDS-1 <sup>ST</sup>	OLDIES-1 <sup>ST</sup>	1 <sup>ST</sup> IN-1 <sup>ST</sup> OUT	FAITH
B-1 <sup>st</sup>					
☞	biker A 1 <sup>st</sup>	✓	✓	✓	*
	biker B 1 <sup>st</sup>	✓	✓	✓	✓

Table 3: ‘faithfulness’ 1-winner solution

difference between two specific candidates. They are not only construction-specific, but their motivation is of a more narrowly technical kind that is typical of faithfulness constraints. The analyses in Table 2 and 3 appear to work literally in the same way: the inputs in Table 3 each specify a property that occurs only in one of the candidates. The same holds for each of the two construction-specific constraints in Table 2. Of course, there are five constraints in Table 2, but the lowest-ranked constraint is superfluous and may even be omitted.

If we do that, we come close to the modelling of German word order optionality by Müller (1999), where we have a constraint subhierarchy in place of a constraint within the main constraint hierarchy, a kind of placeholder “tie break” constraint that can be filled by either A-1<sup>st</sup> or B-1<sup>st</sup>.

We could also formulate the two constraints “A-1<sup>st</sup>” and “B-1<sup>st</sup>” in a way

that comes closer to markedness as usually understood if we described them as forbidding, rather than requiring something, i. e. “ $\neg A-1^{st}$ ” and “ $\neg B-1^{st}$ ”. But this would not change their main function either: giving preference to one specific candidate – it would only do so in a more indirect way.

Again, Table 2 and 3 and the variants that we have just discussed are only different ways of implementing the same underlying logic: adding construction-specific constraints that transform an underspecifying multi-winner system of markedness constraints as in Table 1 into a single-winner system that could be taken to mimick a *generative-enumerative* approach.

I see one conceptual advantage in the faithfulness variant in that it does not add markedness constraints, but one could argue, as I just did, that “ $A-1^{st}$ ” and “ $B-1^{st}$ ” are just faithfulness constraints in disguise.

This may hold for many markedness constraints in the OT syntax literature. For instance, consider a constraint requiring the fronting of a verb in the imperative mood in German imperative clauses,  $FRONTV_{IMP}$ . This constraint simply expresses an aspect of the relevant sentence type in the form of a constraint. If we considered sentence types to be specified in the lexicon, then this constraint would be a faithfulness constraint (because the sentence type would have to be given in the input just like other lexical units), otherwise it would look like a markedness constraint. But it would function in the same way in both cases.

So let us consider some real language examples. Müller (1997) cites the *Ross alcohol rule* for binomials (Ross 1980, “more before less alcohol”, as in *liquor and wine*). What about beverages with an equivalent amount of alcohol? Two such beverages are beer and cider.<sup>3</sup> The prediction of the alcohol rule should be that both possible orders can be found. This is correct, as the following two examples show:

- (1) This **beer-cider** cocktail is perfect for a warm, sunny day  
(<https://thecidercrate.com/the-cherry-berry-orchard/>)
- (2) ... **cider and beer** are a natural fit for some delightful seasonal sipping  
(<https://content.kegworks.com/blog/honey-cider-beer-cocktail>)

So, no addition to Ross’ alcohol rule has to be made, on the contrary: it would be empirically inadequate to posit one constraint each for (1) and (2).

<sup>3</sup>The southern Hessian version of cider, *Ebbelwoi*, is very famous in Frankfurt am Main, Gereon’s first station as a linguist.

Another example in case are attributive uses of adjectives in noun phrases. There are various semantic classes of these, and for some of them natural ordering conditions hold, such that the order (3a) is natural or unmarked, whereas the order (3b) requires specific motivation (e.g. contrastive focus on *blaues* ('blue') also implying givenness for *altes* ('old')).

- (3) a. mein altes blaues Fahrrad  
       my old blue bike  
       b. ??mein blaues altes Fahrrad  
       my blue old bike

With *rostig* ('rusty') in place of *altes* ('old'), the same restriction occurs. However, the order appears to be optional with respect to these two adjectives:

- (4) mein altes rostiges | rostiges altes Fahrrad  
       my old rusty | rusty old bike

If we were to build classes, we could state that colour is a constant property and being old and rusty are dynamic properties that increase over time.<sup>4</sup> Another type of predicates are relational ones like *groß* ('large'), which naturally precede those in (4).

- (5) a. mein großes altes Fahrrad  
       my large old bike  
       b. ??mein altes großes Fahrrad  
       my old large bike

Again, the way we usually proceed is taking optionality as evidence of the absence of a constraint. So the difference between relational, dynamic and constant properties expressed by adjectives is relevant for word order, but the semantic difference between properties like 'old' and 'rusty' is not.

Therefore, we can straightforwardly conclude, our grammar must not contain a constraint relating to this difference. Sadly, this clear line of reasoning has been blurred somewhat in Optimality Theory because one could also assume the relevant constraints to be there, but only "low-ranked".

This relates to the OT way of dealing with language particularities, typological contrasts, language variation and grammatical change. The default OT

---

<sup>4</sup>This, I should add, also holds for animate beings.

answer to all of these issues is constraint reranking. All grammars, by default, are assumed to have the same set of constraints. If a constraint that is active in one language plays no role in another language, then this is not due to absence of the constraint, but its low ranking.

The fundamental problem with this account lies in its ahistoricity. If language change is only constraint reranking and never implies the invention or loss of constraints, then all constraints must have been there from the very beginning of mankind, rooted in general or language-specific cognition.

This may be feasible with respect to constraints that do have a cognitive foundation.<sup>5</sup> But as soon as we discover more subtle cultural aspects to be relevant for word order, like Ross' alcohol constraint, this story becomes completely implausible. Even with respect to the adjectives just discussed, it appears arbitrary that the difference between *rusty* and *old* does not affect word order preferences, while that between *old* and *large* does.

It seems, thus, unavoidable to accept that languages have different constraint sets, not only different rankings of the same set of constraints. Once we accept this, we can go one step further and assume that constraints are only part of one language's grammar if they have observable effects, and then optionality is indeed indicative of the *absence* of constraints.

The conclusion that constraint sets may change over time, that constraints can be introduced or get lost, opens the opportunity to integrate phenomena of grammaticalisation. New grammatical constraints are usually introduced in a very local domain from where the new pattern is then transferred by analogy into further domains, and so it will subsequently become a general constraint of the grammar.

This predicts that a synchronic grammar may display different regulations even for related domains, and of course, there are many candidates for phenomena of this kind. The example I will inspect, finally, are German definite and indefinite pronouns. I have always disagreed (in part) with the assessment put forward e.g. in Müller (1999) that there is a strict ordering condition for definite weak object pronouns in German. The proposal is that (6b,d,f) are equally odd, because they divert from the required NOM < ACC < DAT order:<sup>6</sup>

---

<sup>5</sup>Although I doubt that there are more than very few linguistic constraints that meet this condition.

<sup>6</sup>Please note that the judgements reported in (6) are my own subjective judgements.

- (6) a. dass er sie jemandem vorstellte  
that he.NOM her.ACC someone-DAT introduced
- b. \*dass sie er jemandem vorstellte  
that her.ACC he.NOM someone-DAT introduced
- c. dass ihn ihm jemand vorstellte  
that him.ACC him.DAT someone-NOM introduced
- d. ?dass ihm ihn jemand vorstellte  
that him.DAT him.ACC someone-NOM introduced
- e. dass es ihm jemand vorstellte  
that it.ACC him.DAT someone-NOM introduced
- f. ?dass ihm es jemand vorstellte  
that him.DAT it.ACC someone-NOM introduced

I agree with the assessment reported in the literature only with respect to subject pronouns: the sentence is ungrammatical if the subject pronoun follows another pronoun, see (6b). But the constraint on the relative order of two object pronouns, if there is one, is much milder. Examples of this kind can be found occasionally in corpora,<sup>7</sup> and their acceptability rating appears to be in the range of marked grammaticality.<sup>8</sup> Still, I agree that (6c,e) are the ideal candidates for the unmarked orders, which implies, crucially, that *animacy* is no word order factor for this class of pronouns. Otherwise, (6f) should be preferred over (6e), and (6f) should be rated worse than (6d). There is no evidence, to the best of my knowledge, that would support such a claim.

---

<sup>7</sup>Here is one example from an academic paper:

- (i) Der Streitpunkt besteht darin, dass Turkmenistan aus dem Kanal mehr Wasser, als **ihm es** die Quoten vorschreiben, entnimmt . . .  
“The point of contention is that Turkmenistan takes more water from the canal than the quotas require.”  
<http://pub.ub.uni-potsdam.de/volltexte/2009/2768/>

<sup>8</sup>In one rating experiment, I used the example in (ii) as a filler sentence.

- (ii) Er freute sich, dass er ihm es geschenkt hatte.

Subjects' ratings were chosen between the three levels of unmarked, slightly and strongly marked grammaticality, as well as ungrammaticality. The 88 ratings were distributed in a rather flat manner, with a majority in the range of no or slight markedness (✓ = 21, ? = 29, ?? = 23, \* = 15).



Let us now consider the series of *wh*-pronouns used as indefinite pronouns. In such uses, these pronouns are prosodically weak and may not occupy the clause-initial position in V2-clauses. But clause-internal pronominal clusters are possible. Are they subject to the same ordering constraints as definite pronouns? Not quite. The *animate* variants are subject to the same subject-first condition and optionality with respect to two object pronouns:

- (7) a. Da hat wer                    wen                    begrüßt  
       there has someone.NOM someone.ACC greeted  
       b. \*Da hat wen wer begrüßt
- (8) a. Da hat wer                    wem                    wen  
       there has someone.NOM someone.DAT someone.ACC  
       vorgestellt  
       introduced  
       b. Da hat wer wen wem vorgestellt

But the inanimate *wh*-pronoun is preceded by the animate variants in the unmarked case, no matter which case it has. The alternatives appear to me to be borderline cases of (un)grammaticality.

- (9) a. Da hat wen                    was                    begeistert  
       there has someone.ACC something.NOM fascinated  
       b. ?\*Da hat was wen begeistert
- (10) a. Da hat wer                    wem                    was  
       there has someone.NOM someone.DAT something.ACC  
       gezeigt  
       shown  
       b. ?\*Da hat wer was wem gezeigt

Hence, we have two constraints and the ranking ANIMATE-1<sup>ST</sup> >> SUBJECT-1<sup>ST</sup> for indefinite pronouns, whereas ANIMATE-1<sup>ST</sup> plays a minor role, if any, for the definite pronouns. Just as in the cases discussed above, we identified a semantic factor, animacy, as relevant for word order. But it is only relevant in one of the two related cases.

So both the indefinite and the definite pronouns have their own unique ordering rules. Furthermore, not all languages may have such weak indefinite pronouns. Hence, whatever we find out about these in German may not be

helpful for the understanding of pronouns in other languages. It may be quite an idiosyncratic aspect of the German grammar.

Definite and indefinite pronouns differ in their pragmatics, of course. Definite pronouns are anaphoric and represent discourse topics, indefinite ones are crucially non-anaphoric and introduce new, non-topical and unspecified referents. This pragmatic contrast may be related to the low vs. high relevance of animacy for word order in the two cases,<sup>9</sup> but it certainly does not *enforce* the observed syntactic contrast.

So where are we? I have discussed the claim that optionality is indicative of the absence of markedness constraints and tried to show that this is a reasonable heuristic criterion. It implies, in OT terms, that competitions may have more than one winner. This is a problem from the production perspective: you cannot utter two different sentences at the same time. But how do speakers come to their decisions in those cases? As grammarians we do the right thing when we remain agnostic about this. The contrast between *model-theoretic* and *generative-enumerative*, thus, parallels somewhat that between comprehension and production, or, in behavioural terms, between acceptability and preference. I tend to prefer the comprehension perspective.

This also means, for instance, that corpus data have to be handled with care. Every sentence in a corpus is based on a decision by its speaker/writer in favour of it, and the criteria underlying this decision are a mixed bag: some are linguistically relevant, others aren't, the choice might be accidental. Two variants may be equally well-formed, but still, one may be strongly preferred in language use for non-linguistic reasons. It still has to be clarified, precisely to what extent frequency contrasts are indicative of grammatical factors, for instance, how large such a contrast has to be in order to corroborate the assumption of a grammatical constraint.

So much for my own humble thoughts on these issues. Gereon has spent his whole life on questions like those raised here and could do much better than me in pointing out the intricacies of grammatical frameworks. Thank you very much, Gereon, for your extensive and extremely inspiring contributions to the field. Keep going! Happy birthday!<sup>10</sup>

---

<sup>9</sup> Animacy is of course an organising principle for the paradigm of German *wh*-pronouns, whereas the personal pronouns are differentiated by grammatical gender. But this feature, interesting as it may be, cannot explain either how animacy affects the ordering conditions for these pronouns.

<sup>10</sup> I would like to express my great thanks to the editors of this Festschrift, for taking the

## References

- Heck, Fabian & Gereon Müller. 2013. Extremely local optimization. In Hans Broekhuis & Ralf Vogel (eds.), *Linguistic derivations and filtering – Minimalism and Optimality Theory*, 136–165. Sheffield: Equinox.
- Müller, Gereon. 1997. Beschränkungen für Binomialbildung im Deutschen. Ein Beitrag zur Interaktion von Phraseologie und Grammatik. *Zeitschrift für Sprachwissenschaft* 16(1-2). 5–51. doi:10.1515/zfsw.1997.16.1-2.5.
- Müller, Gereon. 1999. Optimality, markedness, and word order in German. *Linguistics* 37(5). 777–818.
- Müller, Gereon. 2001. Order preservation, parallel movement, and the emergence of the unmarked. In Géraldine Legendre, Jane Grimshaw & Sten Vikner (eds.), *Optimality-theoretic syntax*, 279–314. Cambridge, MA: MIT Press.
- Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind & Language* 28(4). 492–521. doi:10.1111/mila.12029.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retoré (eds.), *Logical aspects of computational linguistics*, 17–43. Berlin, Heidelberg: Springer.
- Ross, John R. 1980. Ikonismus in der Phraseologie. Der Ton macht die Bedeutung. *Zeitschrift für Semiotik* 2. 39–56.

