Sven Kotowski* and Holden Härtl

How real are adjective order constraints? Multiple prenominal adjectives at the grammatical interfaces

https://doi.org/10.1515/ling-2019-0005

Abstract: Adjective order restrictions on attributive adjectives (AORs) have been subject to debate in modern linguistic research for a long time. Most generally, the question whether AORs can be located in grammar as such in rule-based fashion is still unsettled. In the current paper, we largely argue against this view and claim that several of the core data to be explained are preferences based on norms rather than rules. A pragmatic explanation is offered to account for marked or apparently ungrammatical examples. First, we demarcate AORs in the narrow sense against data based on truth-conditional differences, show the sole hard constraint to be found in a distinction between object- and kind-modification, and introduce several of the factors argued to drive AORs in the literature. A large-scale corpus study on German AAN-phrases shows a hierarchy of relative adjectives preceding absolute ones to reliably predict preferences, while temporariness and weight do not. We then illustrate that norm-based preferences can be overwritten via discourse linking and implement markedness in out-of-the-blue contexts pragmatically based on the M-principle. Speculating that AORs in the narrow sense have their origins in more general cognitive principles, our findings support approaches that locate the better part of AORs outside the realm of core grammar.

Keywords: adjective order, temporariness, pragmatics, nominal syntax

***Corresponding author: Sven Kotowski,** Institut für Anglistik und Amerikanistik, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, E-mail: sven.kotowski@uni-duesseldorf.de
Holden Härtl, Institut für Anglistik/Amerikanistik, Universität Kassel, Kurt-Wolters-Str. 5, 34125 Kassel, Germany, E-mail: holden.haertl@uni-kassel.de

1 Introduction

As a classic topic in linguistics, order restrictions on attributive adjectives (henceforth AORs) have received considerable attention in recent years (see, for example, Bouchard 2002; Cinque 2010; Scott 2002). A large variety of factors driving AORs have been proposed in the literature, such as notional and grammatical adjective classes, morphophonological weight, temporariness of the encoded concepts, conceptual iconicity between adjective and noun, and frequency (see, for example, Eichinger 1992; Trost 2006; Wittenberg and Trotzke this issue). However, the nature of AORs as either core syntactic phenomena or merely preferential reflexes of more general cognitive ordering principles is still very much under debate. These two opposed sides of the argument can be exemplified by Scott's (2002: 97) remark that "conjectures as to the psycholinguistic motivation for AOR need not be posed: AOR fall out as a direct consequence of UG" and Bouchard's (2002: 121) claim that while "the classification of properties may interact with language, it is not part of grammar, it does not fall under the object of study of linguistic theory." The examples in (1) and (2) illustrate some of the data to be explained:

(1) a. *mein kleiner grüner Kaktus*
'my little green cactus'
b. ?/**mein grüner kleiner Kaktus*[1]
'my green little cactus'

(2) a. *ein nettes französisches Mädchen*
'a nice French girl'
b. ?/**ein französisches nettes Mädchen*
'a French nice girl'

1 Throughout this paper, asterisks (*) are used to denote ungrammaticality, question marks (?) to denote questionable grammaticality or markedness, and a combination of the two (*/?) markedness tending to ungrammaticality. Note that the examples in (1b) and (2b) pronouncedly gain in acceptability with focus stress on the respective first adjectives, i.e. *mein GRÜNER kleiner Kaktus* and *ein FRANZÖSISCHES nettes Mädchen*, respectively.

Focusing on modified German object nouns in this paper, we argue in favor of some middle ground that informs AORs.[2] Thus, we claim the only hard constraint to be found in the distinction between object- and kind-modification (see also Carlson 2003; Larson 1998), a bipartition that as a default, though not exclusively, coincides with the split between the classes of quality and relational adjectives. In opposition to a variety of approaches that draw on notional adjective classes in their AOR-accounts (see Cinque 1994, Cinque 2010; Scott 2002), we show orders among object-modifying quality adjectives to be preferences based on descriptive norms, i.e. knowledge about the default realization of multiple adjectives derived from generalizations across typical instances (see d'Avis 2013; Härtl 2016). This, in particular, holds for the strong order preference of relative preceding absolute quality adjectives, which turns out statistically significant in a corpus study. Moreover, the factors weight and temporariness, oftentimes claimed informative for AORs (see Cinque 2010; Eichinger 1992; Vendler 1968), are shown to either not be at play at all in our data (weight) or to follow more general order patterns (temporariness). Although norm-based preferences lead to at times strong grammaticality judgements, in particular in out-of-the-blue contexts, we illustrate that contextual enrichment and discourse-embedding can overwrite such preferences. The paper is structured as follows:

Section 2 provides an overview of the adjective classes relevant to our investigation and introduces several of the factors claimed to drive AORs. We demarcate AORs proper from more general order restrictions and disentangle hard constraints and norm-based preferences, both theoretically and empirically. Moreover, we introduce temporariness as a potential predictor for AORs – thus, the chapter at hand also contributes to the ongoing individual-/stage-level debate (see, for example Kratzer 1995 for an overview). Section 3 offers an empirical reflection, reporting on a corpus study on German AAN-phrases including modifiers that encode temporary concepts, picking up several further factors established in the previous section and putting them to the test. In particular, the predictions regarding object- and kind-modification are corroborated, while weight and temporariness are not found to significantly inform adjective order. Finally, Section 4 is more explorative in nature. It introduces several case studies that illustrate the peculiar nature of evaluative adjectives and show discourse-linking phenomena to be informative regarding untypical orders. We offer an application of a pragmatic, manner-based principle to account for the markedness

2 Sections 2 and 3 of this chapter build on Kotowski (2016: Chs. 1/3). Section 2 fleshes out parts of the descriptive and theoretical background initiated there, while Section 3 recaps a corpus study reported on there. Section 4, in turn, picks up a variety of loose ends, offering explanations of preference-based phenomena and approaches to a pragmatic implementation.

of non-preferred orders and conclude the study with considerations of more general cognitive orders as the source of norm-based preferences.

2 AORs as hard constraints and norm-based preferences

This section establishes the constructions and adjective classes relevant to the phenomenon (as well as those that elude AORs) and carves out the differences between what may be called hard constraints and preferences based on norms, i.e. frequency-driven phenomena constituting a norm that allow for counter-examples. The section serves as the basis for Sections 3 and 4.

2.1 Relevant adjective classes

AORs are a phenomenon restricted to the attributive use of adjectives (and largely to prenominal modification in German).[3] Several expressions that do feature in attributive use are outside the scope of this paper: numerals (both cardinal and ordinal), demonstratives, discourse anaphoric/cataphoric expressions (e.g. *abovementioned, following* etc.), and modal or non-subsective adjectives (e.g. *former, future, alleged* etc.). The following provides a typical descriptive typology of the two supra-classes among German adjectives at the heart of AORs, quality and relational adjectives (see, among many others, Levi 1978; Schlücker 2014: 56–57; Trost 2006: Chs. 4 and 5), followed by a brief introduction to a degree-based typology of quality adjectives:

2.1.1 Quality adjectives

Quality adjectives, commonly understood as the *real* property words, display the highest degree of prototypicality, i.e. they tend to occur in all canonical syntactic frames (see Motsch 2002), and the better part of them is subject to grammatical

3 Thus, we will neglect postnominal modification, which is severely restricted in both English and German (unlike in, for example, Romance languages). For languages with postnominal adjective placement as well as (conjectures as to) the relationship between pre- and postnominal placement, see, for example, Bouchard (2002); Cinque (2010); Guisti and Iovino (this issue); Panayidou (this issue).

comparison. Examples for relative and absolute quality adjectives[4] – as well as some of the property concepts they denote – are provided in (3) and (4), respectively:

(3) e.g. evaluatives (*schön* 'beautiful'; *hässlich* 'ugly' etc.), dimension adjectives (*niedrig* 'low'; *eng* 'narrow' etc.), speed adjectives (*schnell* 'fast'; *flink* 'brisk' etc.), or physical state adjectives (*kalt* 'cold'; *heiß* 'hot'; *hell* 'bright' etc.)

(4) e.g. basic color terms (*rot* 'red'; *blau* 'blue' etc.), shape adjectives (*quadratisch* 'square'; *rund* 'round' etc.), defectiveness adjectives (*blind* 'blind'; *stumm* 'mute' etc.), or alternative state adjectives (*lebendig* 'animate'; *tot* 'dead' etc.)

2.1.2 Relational adjectives

Relational adjectives primarily differ from quality adjectives in that they do not modify a noun directly but always fall back on a second point of reference, encoded in the adjective itself. The reference point is usually a second nominal concept (the base of the morphologically complex adjective),[5] the primary syntactic criterion is their restriction to attributive use, and they elude grammatical comparison as well as other forms of gradability (see, for example, Motsch 2004; Schlücker 2014: 56–65; Trost 2006: 124–135). (5) itemizes some of the pertinent subclasses pertaining to general areas of life, which pick out subkinds of the kind denoted by the nominal head, with many authors assigning to them both functional and semantic equivalence with NN-compounds (see, for example, Gunkel and Zifonun 2009: 208–209; McNally and Boleda 2004: 189–190)

(5) e.g. ethnic adjectives (*belgisch* 'Belgian'; *amerikanisch* 'American' etc.), material adjectives (*metallen* 'metal(ic)'; *seiden* 'silk(en)' etc.), or religious

4 See Trost (2006) for a descriptive distinction between relative and absolute quality adjectives that is based on the criterion of grammatical comparison (and not to be confused with the relative-absolute distinction between gradable adjectives; see below). Lack of grammatical comparison for absolute adjectives is claimed to be down to their independence of a subjective or ascertainable comparison value.

5 Some authors also include other reference points into their typologies; for example, temporal, locative, or (text-) deictic ones (see, for example, Eichinger 1992: 319–321; Trost 2006: 140–144). These will be neglected in this chapter.

and cultural terms (*jüdisch* 'Jewish'; *islamisch* 'Islamic'; *sprachlich* 'linguistic' etc.)

2.1.3 Degree-based absolute and relative adjectives

The corpus study discussed in Section 3 employs the degree-based classification system of gradable adjectives introduced in Kennedy and McNally (2005) and Kennedy (2007).[6] The general subdivision along these lines establishes relative gradable adjectives, which are vague in their positive form, and absolute gradable adjectives, which are not, by means of the compatibility of adjectives with degree and proportional modifiers (see also Rotstein and Winter 2004 for a similar typology).

In summary, this paper focuses on the better part of items commonly classified as adjectives, but excludes peripheral subclasses such as modal adjectives, numerals, or discourse anaphoric/cataphoric expressions. The coarse classification contains quality adjectives – including absolute and relative gradable ones in the sense of Kennedy and McNally as well as non-gradable quality adjectives – and those relational adjectives that pick out subkinds of the kinds denoted by the nouns they modify. The corpus study in Section 3 also includes German present and past participles in attributive use.

2.2 AOR-relevant factors and theoretical preliminaries

2.2.1 Modification type and truth conditions

A primary caveat concerns the difference usually captured by the two notions of 'parallel' and 'hierarchical modification' (see, for example, Cinque 1994: 101–103; Scott 2002: 92; Sproat and Shih 1988: 477–478). Parallel modification as in (6) applies to structures in which the adjectives appear to modify the respective nominal head independently of each other.[7] In turn, in hierarchical modification

6 While oftentimes applied in theories on AORs, the set-theoretic classification into intersective, subsective, and non-subsective modifiers (see Kamp 1975; Partee 1995) is highly problematic and, in fact, ill-suited for modeling AORs. In particular, noun-independence is a dubious concept (cf. Bouchard 2002: 62–63) and most formal semanticists working on adjective typologies in fact aim at modelling the better part of the subsective quality adjectives in intersective fashion; see Kotowski (2016: Chs. 1.2.2.2, 3.2.2.3) for discussion.

7 Parallel modification is accompanied by either comma intonation and/or overt conjunctions in between the modifiers.

as in (7), an adjective modifies the complex set established by (multiple) adjective-noun-clusters. In this case, the modifier in farther distance from the head scopes over the ones in closer proximity to the head, specifying reference in recursive fashion:[8]

(6) PARALLEL MODIFICATION Schema: [A + A + A + N]
   a. *John loves Fury, his* [*cute, beautiful, tall riding horse*].
   b. *John loves Fury, his tall, beautiful, cute riding horse.*

(7) HIERARCHICAL MODIFICATION Schema: [A + [A + [A + N]]]
   a. *John went to a horse breeder looking for a* [*cute* [*beautiful* [*tall riding horse*]]].[9]
   b. [?]*John went to a horse breeder looking for a tall beautiful cute riding horse.*

The differences in acceptability in (6a/b) and (7a/b) appear to be down to comma intonation, and thus to the distinction between the two modification types. Although the corpus study in Section 3 cannot do full justice to the complexity of the phenomenon, working with written corpora, these considerations lead to the heuristic exclusion of all structures with commas in between adjectives from the analysis.[10]

Moreover, AORs in the narrow sense only apply if the reversal of adjective orders is degraded in acceptability and at the same time does not alter the truth-conditions of the entire DP if adjective senses are kept constant (see Champollion 2006: 2–3; Svenonius 1994: 450–451; Teodorescu 2006; Zifonun et al. 1997: 1992–1994). In contrast, the examples in (8) show two possible orders that appear equally acceptable – however, different truth conditions apply to the reversed orders:[11]

8 'Recursivity' is used as a semantic notion here. Depending on the respective syntactic model, hierarchical modification may well be represented via stacking of functional projections of different kinds, rendering it recursive only in a general syntactic but not a specific sense (see Kotowski and Härtl 2011; Van der Hulst 2010 for the difference between specific and general recursion).

9 It is unclear in how far hierarchical modification presupposes restrictive modification (see Zifonun et al. 1997 in favor of this argument).

10 Moreover, a further well-known effect that reliably overrides AORs concerns fronted constructions with focal stress (see, for example, Scott 2002: 92–93; Teodorescu 2006: 400) as in **a red beautiful car* opposed to *a RED beautiful car*. This could not be tested in the corpus study.

11 As one anonymous reviewer rightly points out, truth-conditional equivalence is difficult to operationalize. Throughout the examples we discuss as well as in the studies we report on, we rely on subjective judgements and at times consulted informants regarding potentially differing interpretations of AAN-phrases and reversed adjective orders.

(8) a. *the counterfeit new banknote*
a.' *the new counterfeit banknote*
b. *a dangerous dead animal* (example adopted from Svenonius 1994: 451)
b.' *a dead dangerous animal*

Thus, we are either dealing with a counterfeit exemplar of a recently issued banknote in (8a) or some new version of a counterfeit bill in (8a'), while (8b) denotes an animal that is likely dangerous due to being dead – for example, a decaying corpse spreading germs – and (8b') some dangerous animal – for example, a beast of prey – that happens to be dead.[12]

2.2.2 Modificational layers and notional classes

Many authors agree that some form of pluricentric or multilayered configuration is best able to capture AORs. The consensus is that the orders are driven semantically and/or pragmatically, with, for example, quantifying modification occuring in farther distance from the head noun than qualifying modification, which in turn occurs in farther distance than classifying modification (see Eichinger 1992; Rijkhoff 2002: Ch. 7, Rijkhoff 2010; Ten Wolde this issue). This is, for example, reflected in Eichinger's (1992: 327) modification zones, adapted here as (9):

(9) EICHINGER'S MODEL OF 3 MODIFICATION ZONES

article	**article classifiers**		**qualifiers**		**noun classifiers**		noun
	quanti-fiers	referential adjectives	evalua-tives	qualita-tives	descrip-tives	classi-fiers	
	referential/situating		**qualifying**		**classifying**		

Commonly, however, AORs among quality adjectives are treated as hierarchies of notional property classes, either in terms of preferences or labile ordering constraints (see, for example, Bache and Davidsen-Nielsen 1997: 462–472; Dixon 1982; Payne and Huddleston 2002: 453–454) or as hardwired constraints directly written into the syntax and – given their cross-linguistic robustness –

12 An explanation for these reading differences can be found in some form of a head primacy principle, stating that in modifier-head-structures, the head has interpretative primacy and is interpreted relative to the context created by the whole constituent, while modifiers are interpreted relative to the local context created from the former context by the interpretation of the head (see Partee 1995: 334).

presumably a matter of UG. A particularly fine-grained example of such a cline is illustrated in Scott (2002: 114), adapted here as (10):[13]

(10) SCOTT'S UNIVERSAL HIERARCHY NOTIONAL ADJECTIVE CLASSES
(Determiner > Ordinal Number > Cardinal Number >) Subjective Comment > ?Evidential > Size > Length > Height > Speed > ?Depth > Width > Weight > Temperature > ?Wetness > Age > Shape > Color > (Nationality/Origin > Material > Compound Element)

Yet, cartographic approaches that assume concrete functional projections based on notional classes (cf. Cinque 1994, Cinque 2010; Laenzlinger 2005; Ramaglia 2011; Scott 2002; Sproat and Shih 1988) have proven unsustainable, primarily due to empirical undergeneration (cf. Kotowski 2016: Ch. 3.2; Svenonius 2008: 34–39; Truswell 2009: 526–528). For example, Scott's (2002) cline in (10) above predicts UG-driven orders of subjective comment (i.e. evaluation) preceding size adjectives and shape preceding color. The following German examples in (11) and (12) are extracted from the DeReKo-corpus, show the reversed orders, and are fully acceptable without applying focal stress or comma intonation.

(11) SIZE > > EVALUATION
Alle Zahlen sind in den Geburtsdaten von ihr und ihrer Familie enthalten. Mit einem größeren Gewinn würde sie sich einen großen schönen Audi kaufen.
'All numbers are included in her and her family's birthdays. With a larger payout she would buy a big beautiful Audi.'
(NON08/SEP.04230 Niederösterreichische Nachrichten, 08.09.2008, NÖN Großformat S. 44; LOTTO&p; JOKER-ZIEHUNGSERGEBNISSE VOM 7. SEPTEMBER)

(12) COLOR > > SHAPE
[...] *eine kreisrunde Kupferplatte mit dem Durchmesser 5 cm, in die drei rote runde Glasscherben und drei weiße kreisrunde Perlmuttscheiben eingelassen sind* [...]
'[...] a circular copper plate, 5cm in diameter, into which three red round pieces of broken glass and three white circular nacre discs are set in [...].'
(WPD/HHH.07860 Christoph73; Topinambur; 1: Holle, In: Wikipedia - URL: http://de.wikipedia.org: Wikipedia, 2005)

13 The brackets are our modification and only the non-parenthesized classes are quality adjectives. Question marks are Scott's, indicating his uncertainty of the necessity of the respective notional classes.

An equally severe caveat concerns the available projections: it is unclear how any of the proposed hierarchies could accommodate all possible quality adjectives. For example, Scott's is already fine-grained and therefore undergenerates, yet seems unable to capture even highly frequent adjectives such as *open, closed, married*, or *empty*. While notional classes should not easily be dismissed as factors in predicting AORs, they are thus outright untenable if conceived of as core syntactic phenomena that can only be circumvented via peculiar stress patterns.

2.2.2.1 General basis of hierarchies

More globally, hierarchies such as the one in (10) are claimed to reflect universal tendencies to realize modifiers along the lines of a variety of gradual ordering principles. First, (more) subjective modifiers are claimed to precede more objective ones (see Hetzron 1978; Scontras et al. 2017) – this tendency is commonly claimed to show a grammatical reflex in that notional hierarchies can be grouped into clines of relative adjectives preceding absolute ones (see, for example, Laenzlinger 2005; Truswell 2009; Wittenberg and Trotzke this issue).[14] Second, modifiers applicable to a larger number of different nouns are hypothesized to precede less applicable ones (see Champollion 2006; Seiler 1978). Finally, modifiers encoding (more) temporary concepts are often assumed to precede modifiers encoding (more) permanent ones. While the applicability criterion is largely neglected here, the relative-absolute divide and the temporariness criteria will be at the heart of the corpus study in Section 3 (for the latter see Section 2.2.3 below).

Arguably, the alleged constraints of (more) subjective preceding (more) objective and of temporary preceding permanent modifiers also show up in the well-known behavior of relational adjectives and their realization close to the modified noun (cf. among others Rijkhoff 2010: Ch. 7; Schlücker 2014: 60–61). This behavior is standardly explained with the inherent classificatory function of relational adjectives in combination with either modification zones – the classifying zone realized closer to the noun than the qualifying one (see above) – or some form of DP-partitioning (cf., for example, Arsenijević et al. 2014; Vergnaud and Zubizarreta 1992; see also Kotowski 2016: Chs. 3/4 for an

14 Or into subsective and intersective adjectives – a distinction which has been argued to not be informative as regards AORs; see Section 2.1 above. Note that we remain agnostic as to whether relative and absolute can easily be mapped to subjectivity – for example, evaluatives are usually claimed to be more subjective than dimension adjectives (both relative), while certain absolute adjectives are argued to allow for greater interspeaker uncertainty than certain relative ones (see Solt 2016 for discussion).

implementation). In this vein, relational adjectives are the prototypical kind- or type-modifiers, while quality adjectives are the typical object- or token-modifiers. The picture is complicated by the fact that quality adjectives in certain collocations can also function as kind-modifiers and relational adjectives oftentimes develop quality polysemes over time. However, as shown in (13), their behavior will then follow the one typical for the respective modificational function (cf. Schlücker 2014; Trost 2006).

(13) a. **an urban nice park*
b. **the jersey is yellow*[15]
c. **the beautiful and yellow jersey*
d. *the cake was heavenly*

Thus, relational adjectives do not occur outside quality adjectives as in (13a), quality adjectives functioning as kind-modifiers do not feature predicatively as in (13b), object- and kind-modifiers cannot be coordinated as in (13c), while quality polysemes that developed out of relational adjectives can, for example, be used predicatively as in (13d). The corpus study reported on in Section 3 makes use of relational adjectives as one of its output classification types and, unlike for possible AORs between quality adjectives, the split between kind- and object-modifiers is hypothesized to be a core grammatical constraint.

2.2.3 Temporariness and weight

2.2.3.1 Temporariness

Across linguistic traditions and frameworks, temporariness is considered a further factor fueling AORs (see, for example, Cinque 2010: 6–7; Eichinger 1992: 321–322; Eroms 2000: 270–271; Halliday 2014: 381; Larson 1998; Posner 1980: 71–72). The general claim in this regard can be summarized as follows: adjectives – and adnominal modifiers in general – that encode (more) temporary property concepts occur farther from the head noun they modify than adjectives encoding (more) enduring properties. In descriptive and functional frameworks, temporariness is usually not formalized, but merely stated as a continuum, and some examples the pertinent literature explicitly argues to be driven by the temporary > > permanent hierarchy are listed in (14).

15 The judgements for (13b–d) are obviously bound to specific interpretations: *yellow jersey* as the leader's jersey in the Tour de France, i.e. as a kind-modifier, and *heavenly* as an evaluative quality adjective.

(14) a. *a new red ball* (adopted from Halliday 2014: 381)
b. *einem betrunkenen jungen Mann* (adopted from Eichinger 1992: 322)
'a drunk young man'
c. *die geöffneten umrankten Fenster* (adopted from Eroms 2000: 271)
'the opened entwined windows'

The examples in (14) are all cases of two quality adjectives, i.e. the claim entails that it is among this class that temporariness predicts word order. In formal semantics, the temporariness of predicate classes has been modelled as the difference between individual- (IL) and stage-level-expressions (SL), which is argued to cut at least through the verb and adjective classes (for theoretical background and pertinent constructions, see Carlson 1980; Fernald 2000; Kratzer 1995). While this notion of a grammatically encoded, fundamental split among word classes is far from uncontroversial (see, for example, the discussions in Jäger 2001; Maienborn 2003), the distinction is also applied to account for very specific AORs. Thus, for example, Cinque (2010) and Larson (1998) both argue for configurations in which SL-modifiers precede IL-modifiers on the basis of adjectival doublets as in *The visible*$_{SL}$ *visible*$_{IL}$ *stars include Capella*. (from Larson 1998: 155), in which the only conceivable readings assign SL-status to the outer and IL-status to the inner modifier. Leaving aside their unnaturalness, such data can be argued to spread over two modificational layers and therefore do not relate to the phenomenon of AORs among quality modifiers. It is the claim that within the layer of quality modification temporariness predicts orders that is empirically put to the test in the corpus study in Section 3.

2.2.3.2 Weight

A further, ceteris-paribus factor often used to explain AORs is morphophonological weight (see, for example, Behaghel's (1909) and Pāṇini's laws;[16] see Cooper and Ross 1975). Following the general tendency that – everything else being equal – morpho- and phonologically longer or weightier constituents or phrases tend to follow shorter ones in coordinated as well as stacked constructions, the claim essentially entails that if two adjectives are in principle interchangeable, the weightier adjective follows the less weighty ones. In German as well as English prenominal adjective sequences, then, weight is argued to function as a predictor for preferred sequences, in that weightier adjectives are encoded in closer proximity to the head noun than less weighty ones (see, among others,

16 Cf. Behaghel's *Gesetz der wachsenden Glieder*, i.e. the 'law of increasing terms/constituents'.

Eichinger 1992: 324–325; Rijkhoff 2002: Ch. 7; Vendler 1968: 121–122). Weight will be another factor investigated into in the corpus study below.

2.3 Summary

Restricting the scope of investigation to quality and relational adjectives, the only hard constraint the data appear to allow for is the distinction between object- and kind-modification. The latter modification pattern includes relational adjectives as defined above as well as quality adjectives functioning as kind-modifiers, as they, for example, materialize in lexicalized AN-expressions and phrasal names. With respect to quality adjectives as object-modifiers, we propose to regard all other factors introduced in this section – notional class, the relative-absolute distinction, temporariness, and weight – as fueling norm-based preferences, not constituting core grammatical phenomena, on the basis of the discussed data. The corpus study reported on in the following section puts several of these predictions to the test, primarily with respect to the order of quality adjectives as object-modifiers and relational adjectives, the relative-absolute distinction, temporariness, and weight.

3 Corpus study

3.1 Rationale

Kotowski (2016) reports on a large-scale corpus study that makes use of the different assumptions and conjectures in the realm of AORs introduced in Section 2. The study hypothesizes that multiple prenominal adjectives in German in hierarchical modification are ordered at least preferentially and therefore puts several factors to the test: Searching the DeReKo for AAN-phrases in both of their two possible orders, it operationalizes seven different output classes, includes temporariness as a factor, and analyzes the output with respect to the weight criterion. While we are aware of the intricacies of reading off grammatical constraints from corpus searches, the minimum assumption is to find significant differences between arbitrary distributions, preferential orders (for example, relative preceding absolute adjectives), and supposedly core grammatical distinctions that bear on word order (here, relational adjectives being realized closer to the head noun than quality adjectives).

3.2 Procedure, items and hypotheses

One of the word-class tagged subcorpora (the TAGGED-T archive)[17] of the German Reference Corpus DeReKo was searched via the Cosmas II web interface for AAN-phrases. The ten predefined adjectives in (15) were used for a total of 20 searches.

(15) INPUT ADJECTIVES
betrunken 'drunk', *dreckig* 'dirty', *hungrig* 'hungry', *leer* 'empty', *leise* 'quiet', *müde* 'tired', *nackt* 'naked', *nass* 'wet', *wütend* 'angry', and *zufrieden* 'content'

All of the input adjectives denote temporary properties, both intuitively and in the sense of the IL-SL-distinction. Passing several of the established tests for SL-hood (see Fernald 2000; Kratzer 1995), they were classified as both 'SL' and 'absolute' in the study.[18] In order to test the preferred placement of temporary adjectives, 20 queries were run in total; two for each input adjective, once as the first of two prenominal adjectives ($A_xA_{att}CN$) and once as the second ($A_{att}A_xCN$), where 'A_x' stands for the predefined adjective, 'A_{att}' for an arbitrary attributive adjective, and 'CN' for an arbitrary common noun. The output adjectives – the respective second adjectives searched for via the part-of-speech tag 'MORPH(ADJ at)' – were assigned to the seven grammatical classes in (16) and (17).[19]

17 TAGGED-T contains roughly 40% of DeReKo's main archive with publication dates up to 2009, i.e. about 1.5 billion word tokens spread over 26 corpora, mostly from newspapers, press agencies, and internet sources. It allows searching for word-class annotated items in certain positions, e.g. 'attributive adjective' or 'common noun', and was thus crucial for the study at hand.

18 All "SL-adjectives" are absolute (see Kennedy and McNally 2005; Kennedy 2007). The sole criterion for the input adjectives was that they pass several tests for SL-hood. As one anonymous reviewer remarks, the results may be somewhat skewed due to their preference for nouns denoting animate beings. Although at least four out of the ten input adjectives are fully compatible with nouns denoting inanimate objects, we will come back to this point in the discussion in Section 3.4.

19 The subdivision into permanent and temporary classes is in fact artificial from a grammatical standpoint, but was necessary for classifying the output according to the research objectives.

(16) PERMANENT OUTPUT CLASSES

1. RELATIVE PERMANENT ADJECTIVES (REL_PERM)
   e.g. evaluatives (*schön* 'beautiful'), dimension adjectives (*groß* 'big')
2. ABSOLUTE PERMANENT ADJECTIVES (ABS_PERM)
   e.g. color terms (*schwarz* 'black'), shape adjectives (*rund* 'round')
3. NON-GRADABLE PERMANENT ADJECTIVES (NONGRAD_PERM)
   mostly past participles (*verheiratet* 'married')
4. RELATIONAL ADJECTIVES (RELATION)
   e.g. ethnic (*italienisch* 'Italian') and material adjectives (*hölzern* 'wooden')

(17) TEMPORARY OUTPUT CLASSES

5. TEMPORARY ADJECTIVES (TEMP)
   the class of the input adjectives
6. NON-GRADABLE TEMPORARY ADJECTIVES (NONGRAD_TEMP)
   mostly temporary past participles (*gefesselt* 'chained)
7. PRESENT PARTICIPLE (PRES_PART)
   all present participles that allow for temporary readings and pass SL-hood tests (*spielend* 'playing')

The output-pairs, viz., all combinations of input- and output-adjectives, were coded for three different values of morphophonological weight: "-1" = the first of two adjectives has a higher weight (i.e. more syllables) than the second; "0" = the two adjectives are equal in weight; "1" = the second adjective has a higher weight. The hypotheses below all derive from adjective classes and temporariness, while weight was tested independently for all output classes.

The following hypotheses were formulated: first, HYPO (null hypothesis) – predicting no statistically significant difference between the seven output classes and the way they cluster with the input adjectives. Second, HYP1 (SL > > IL) – if assumptions regarding the SL > > IL claim are correct, we expect the permanent classes 1.–4. to be found robustly closer to the head noun than the input adjectives and to significantly differ in this respect from the temporary classes 5.–7. Third, HYP2 (RELATIVE > > ABSOLUTE) – if assumptions regarding the relative > > absolute hierarchy are correct, we expect class 1 (REL_PERM) to be more reliably found preceding the input adjectives than classes 2.–7. Finally, following the core division between object- and kind-modification outlined in Section 2.2.2 above, HYP3 predicts relational adjectives as kind-modifiers to be found closer to the noun than the input adjectives.

3.3 Results

Searching the corpus via the 20 queries yielded an overall output of 6.622 hits. Manual cleansing led to a remaining sample size of $N = 508$.[20] Table 1 provides the overall occurrences of the seven subclasses of the respective output adjectives and their distribution given the respective input adjectives as either the first (A_1) or the second (A_2) of two prenominal adjectives, while Figure 1 illustrates the overall distributions by percentage for A_1 and A_2, respectively.

Table 1: Absolute occurrences of output adjective classes depending on the position of the input adjectives. Figures include the total 508 hits for all 20 search queries.

input position output class	A_1	A_2
REL_PERM	56	102
ABS_PERM	42	11
NONGRAD_PERM	36	30
RELATION	147	0
TEMP	11	13
NONGRAD_TEMP	11	20
PRES_PART	11	18
total	314	194

The relationship between the position of the input adjective and the output class was significant; χ^2 (6, $N = 508$) = 164.36, $p < 0.001$. The classes RELATION and ABS_PERM reached significance, while REL_PERM ($A_2 = 64.6\%$) showed at least a clear trend. Being a constant, RELATION was excluded from further analysis, reducing the sample size to $N = 361$.[21]

A multinomial logistic regression was performed to compare data from the dependent variable's different categories (output adjectives) and their behavior

20 Cleansing was necessary and occurred for a large variety of reasons, most prominently commas in between the two adjectives: This was necessary as a heuristic to capture the difference between parallel and hierarchical modification; see Section 2.2.1. Other reasons for exclusion were clear-cut idiomatic expressions or lexicalized phrases (e.g. *interessante nackte Tatsachen* 'interesting hard facts', lit. interesting naked facts, was excluded due to the lexicalization of *nackte Tatsachen*), and wrong part-of-speech tags.

21 The relation between input adjective position and output class remained significant; χ^2 (5, $N = 361$) = 34.71, $p < 0.001$.

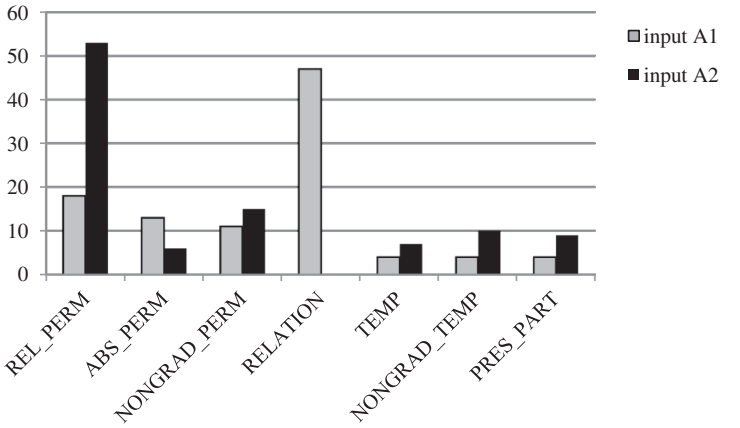

Figure 1: Distribution by percentage of output adjective classes given the input adjective as the first (gray bars) or the second adjective (black).

as regards the independent variable (adjective position) against the reference category TEMP.[22] Figure 2 visualizes the results.

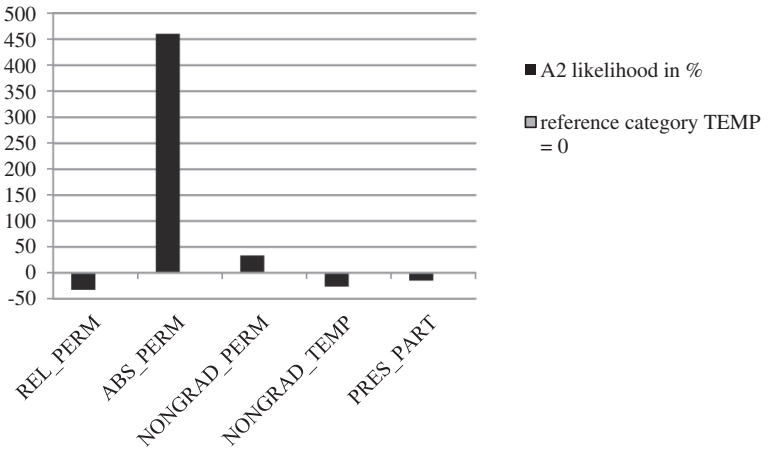

Figure 2: Odds ratios (likelihood) of output adjectives from the different classes to feature as the second of two adjectives in comparison to the reference class TEMP. The reference category is consequently set to '0'.

22 TEMP was chosen as the reference category because the input adjectives were classified as TEMP themselves and, consequently, the found distribution of A_1s and A_2s is arbitrary for combinations of input adjectives and TEMP adjectives (see Table 1 and Figure 1).

Only ABS_PERM reached statistical significance ($p = 0.002$) – if the input adjective is in position A_1, it is roughly 460% more likely, compared to the reference category TEMP, that the output adjective in position A_2 is of category ABS_PERM (Exp(B) = 5.604). In contrast, Exp(B)s cluster around the reference value 1 for the other four categories (REL_PERM, NONGRAD_PERM, NONGRAD_TEMP, PRES_PART) and none of them reached significance (but see the discussion below for relative adjectives).

Finally, an ANOVA calculated on the influence of output class in combination with input position on the independent variable WEIGHT[23] was significant; $F(11, 298) = 36.68$, $p < 0.001$. Somewhat counterintuitive, these results show weight to not predict orders; see the discussion section below. Table 2 itemizes the post-hoc tests (Tukey HSD), while Figure 3 illustrates the weight measure distribution across output classes and input position.

Table 2: Post-hoc results (Tukey HSD) of the ANOVA for the intra-output-category distribution of adjectives depending on adjective weight.

input position output class	A_1	A_2	A_1 vs A_2 mean diff. / significance
REL_PERM	M = −0.50, SD = 0.51	M = 0.00, SD = 0.65	diff = 0.50, $p < 0.001$
ABS_PERM	M = 0.03, SD = 0.28	M = −0.70, SD = 0.48	diff = 0.73, $p = 0.002$
NONGRAD_PERM	M = 0.83, SD = 0.38	M = −0.90, SD = 0.31	diff = 1.72, $p < 0.001$
TEMP	M = 0.17, SD = 0.41	M = −0.15, SD = 0.56	diff = 0.32, $p = 0.974$
NONGRAD_TEMP	M = 0.91, SD = 0.30	M = −0.94, SD = 0.24	diff = 1.85, $p < 0.001$
PRES_PART	M = 0.75, SD = 0.71	M = −0.94, SD = 0.24	diff = 1.69, $p < 0.001$

3.4 Discussion

The null hypothesis (viz., no correlation between input adjective position and output class) is rejected, given the trends for the output classes REL_PERM, ABS_PERM, and RELATION (the latter two significant). The theoretical assumptions outlined in Section 2 are supported, as relational adjectives as kind-modifiers are the only class that introduces a hard constraint: As a statistical constant, they are always realized closer to the noun than the input adjectives (upon prior exclusion of lexicalized phrases such as *nackte Tatsachen* 'hard facts'). Thus, HYP3 (input adjectives > > RELATION) is corroborated. All other classes allow for exceptions to statistical patterns (for example, input

23 In fact, the ANOVA was conducted for only 5 out of the 10 input adjectives (*betrunken, hungrig, leer, nackt*, and *nass*), as the other 5 adjectives did not yield significant results in isolation. This led to a final minor adjustment in sample size ($N = 310$ as opposed to $N = 361$).

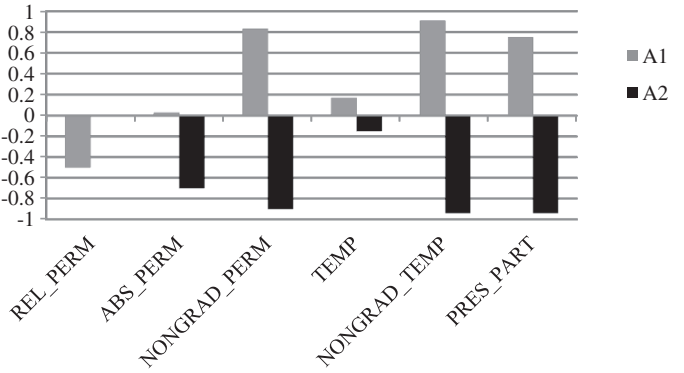

Figure 3: Measures of morphophonological weight for the respective two adjectives across output classes with the input adjective either as the first (gray bars) or the second adjective (black).

adjectives > > REL_PERM or ABS_PERM > > input adjectives). HYP1 (SL > > IL) is also rejected: REL_PERM and ABS_PERM, the two major permanent classes besides the eliminated class RELATION, show inverse behaviors with respect to input adjective positions, the former predominantly featuring as the first, the latter as the second of two prenominal adjectives. Thus, temporary adjectives by no means generally precede permanent ones, but, on the contrary, tend to follow relative ones (being absolute themselves).

Logically following from the reasons for rejecting HYP1, HYP2 (relative > > absolute) is corroborated: Relative adjectives can be said to precede absolute ones with statistical normality (see Schurz 2001 for the notion of statistical normality). Thus, they precede the absolute input adjectives and, by transitive reasoning, can be argued to also precede absolute ones in general. Post-hoc analysis allows for identifying several patterns. Upon exclusion of age adjectives that modify nouns denoting animate beings, the class REL_PERM ($p = 0.036$) also reaches statistical significance, supporting claims that age adjectives form clusters of conceptual iconicity with certain nouns. While age adjectives are clearly relative, they have been argued to frequently feature relatively close to nominal heads denoting animate beings and thereby perform a quasi-classificatory function (cf. Eichinger 1992: 322–323), with several authors drawing on Behaghel's (1909) first law[24] of semantic iconicity (Bouchard 2002: 120–121, Bouchard 2011; Eichinger 1992; Trost 2006: Chs. 21/22). Thus, these peculiar results may well be partly down to the input adjectives used and their preference for a certain

24 Cf. his first law (*Erstes Behaghelsches Gesetz*), which roughly says that the conceptually closely related also occurs in close proximity.

subclass of nouns. Arguably for the same reason, no shape adjectives feature in the ABS_PERM output, with color terms by far the largest group of items in this class (42 out of 53 adjectives) – while this latter point appears accidental, it speaks in favor of a distinct behavior of the notional class color that preferentially merges very low close to the head noun. Yet, again, this is clearly not a rule that would not allow for exceptions, as the following example from our data shows:

(18) *Rund 20 Jahre alt, etwa 1,75 Meter groß, sehr schlank, schwarze Cordhose, schwarze dreckige Schuhe, weißes T-Shirt, kariertes Hemd* [...]
'Around 20 years old, ca. 1.75m tall, very slender, black corduroys, black dirty shoes, white t-shirt, checkered shirt [...]'
(RHZ97/OKT.00313 Rhein-Zeitung, 01.10.1997; prostituierte mt)

Finally, while the factor weight does not predict adjective order in the study at hand (note that it may well contribute in other studies that balance their input adjectives regarding weight and include more adjectives), it is nonetheless instructive. In general, with input adjectives being short themselves, weightier expressions tend to follow A_1-input-adjectives and to precede A_2-input-adjectives. While this holds true for all output classes, it is tellingly most pronounced for the non-significant classes (NONGRAD_PERM, TEMP, NONGRAD_TEMP, and PRES_PART). This observation corroborates the nature of the REL_PERM and ABS_PERM classes as significant predictors for preferred adjective orders among quality adjectives. It is also worth noting that the morphologically complex relational adjectives (RELATION) had already been excluded from the data and did not feature in the weight measures. To sum up, the corpus study corroborates the reasoning that the distinction between object- and kind-modification is the sole hard constraint, shows the relative > > absolute hierarchy to significantly predict adjective order among object-modifying quality adjectives, and rejects both temporariness and weight as significant factors.

4 AORs and the pragmatics interface

Building on the findings from the above sections, we have arrived at the conclusion that a cartographic implementation of fine-grained AORs for quality adjectives is untenable. AORs of this type are not hardwired in the grammar of a language. Instead, they appear to be rooted in preferences for orders in specific adjective combinations and regularities grounded on norms, i.e. in

generalizations across occurrences of multiple prenominal adjectives.[25] An important property of knowledge about norms is that it has to accept exceptions. An exception can occur on a regular or on a non-regular basis.[26] In the case of multiple adjectives, this view entails that linguistically acceptable reversals of typical orders occur consistently, albeit with a lower frequency. In the following, we report on data from a case study in support of this view, under the assumption that an analysis of AORs as norm-based preferences applies to semantically identical order variations alone. In particular, we will show discourse configurations to be a factor which can influence adjective orders. The markedness perceived with reversed orders in out-of-the-blue contexts will then be attributed to a manner-based implicature that cannot be resolved in environments of this sort.

4.1 Case study on untypical adjective orders

The linguistic acceptability of certain reversals of typical adjective orders is confirmed by the results of an exploratory case study, in which we examined the numerical nature of adjective order preferences for a selection of individual quality adjective combinations. For example, for the combination of the dimension adjective *klein* 'small' and the color adjective *schwarz* 'black', our search[27] in DeReKo reveals a strong preference for *klein* > > *schwarz* ($N = 316$) over the reversed order *schwarz* > > *klein* ($N = 5$). An order preference can also be observed for the combination of the shape adjective *rund* 'round' and the color adjective *weiß* 'white', for which the query returned 16 hits for the order *rund* > > *weiß* and only 5 hits for the reversed order. Qualitative inspection reveals that, with these two combinations, no systematic difference in meaning can be detected for the two orders. Consider the following two examples for illustration.

25 With this statistical understanding, norm and convention are not identical notions as we assume conventions to represent standards, which are accepted as such and follow from the weaker notion of norm.

26 For example, a male duck does not lay eggs and is a predictable exception from the norm that ducks lay eggs. In contrast, an adult duck which cannot fly is presumably injured and is, thus, no regular exception from the norm that ducks can fly.

27 Invalid hits (i.e. combinations with commas in between adjectives, comparatives, and adjective-modifying etc.) were excluded from the analysis. Also, adjective combinations were balanced with respect to weight (same number of syllables) and general frequency (at maximum, they were two frequency classes apart in the *Leipzig Wortschatz*-corpus; see < < http://wortschatz.uni-leipzig.de/ > >).

(19) Typical order: *klein* > > *schwarz* (*N* = 316)
Ideen und kleine Geschichten sammelt Garm, der sich momentan in der Freistellungsphase der Altersteilzeit befindet, das ganze Jahr über. Sein Markenzeichen ist ein kleiner schwarzer Hut.
'Ideas and short stories are collected by Garm, who is currently in the release period of partial retirement, throughout the year. His trademark is a little black hat.'
(RHZ08/JAN.23351 Rhein-Zeitung, 28.01.2008; Garm: Seit 41 Jahren in der Bütt)

(20) Untypical order: *schwarz* > > *klein* (*N* = 5)
Alle drei Monate kommt die Familie mit ihrer Ernte zum legalen Kokamarkt von La Paz; vier Säcke sind es pro Fahrt, schätzt die Frau mit dem schwarzen kleinen Hut.
'Every three months, the family comes to the legal coca market in La Paz; four bags per trip, as estimated by the woman with the black little hat.'
(RHZ06/APR.20647 Rhein-Zeitung, 24.04.2006; Morales im Konflikt mit Kokabauern)

In the above examples, a reversal of the preferred adjective order does not entail a semantic difference[28] and, importantly, both orders are perceived as unmarked, thus fueling the view that the ratio of the two adjective orders is based on a norm. The notion of norm cannot itself be used to explain the occurrences of an untypical order, rather, we assume it to be a measure of the different order distributions. Further, a norm-based reasoning on order ratios cannot be applied to combinations in which the ordering has a specific semantic rationale. One example is the combination of the evaluative adjective *schön* 'beautiful' and the dimension adjective *groß* 'big', for which our corpus search returned 127 hits for *schön* > > *groß* and 13 hits for the reversed order. Here, however, instances of the order *schön* > > *groß* include several cases in which a causal connection between the two modifiers can be assumed to hold, as illustrated in the following examples.

28 We remind the reader that AORs only apply if the reversal of the adjective order is degraded in acceptability and, at the same time, does not alter the truth-conditional content of the construction, see Section 2.2.1.

(21) *schön* > > *groß* (*N* = 127)

a. *Die meisten Möbel wurden pünktlich geliefert, nur das schöne große Bett nicht. Ich liege also immer noch auf der Matratze, die ich in der Mansarde entdeckt hatte.*
'Most of the furniture were delivered on time, but not the beautiful large bed. So, I still lie on the mattress, which I had discovered in the attic.'
(RHZ09/FEB.15301 Rhein-Zeitung, 18.02.2009; Schließlich trennen wir uns ...)

b. *Jetzt haben wir schöne grosse Fenster; dort, wo wir früher wohnten, war es immer so düster.*
'Now we have beautiful large windows; where we lived before it was always so gloomy.'
(A98/SEP.54881 St. Galler Tagblatt, 05.09.1998, Ressort: TB-SG (Abk.); getroffen)

In these examples, the interpretation of the DPs in question supports the implication that the respective referents, i.e. the bed and the windows, are evaluated as beautiful as a consequence of their being large. As mentioned above, this reading is not the only one available and we do not claim that the two possible interpretations can easily be disentangled.[29] Crucially, however, an analogous interpretation is not manifest in the reversed order. Consider the following example:

(22) *groß* > > *schön* (*N* = 13)
Die Mitglieder dieser neuen Bewegungen hingegen gehen von Haus zu Haus und fragen, wie es den Leuten geht. Und dann stellen sie eine große schöne Kirche hin und laden zum Singen und zum Beten ein, versprechen eine bessere Zukunft.
'The members of this new movement, by contrast, go from house to house and ask how the people are doing. And then they erect a big beautiful church and invite to sing and pray, promise a better future.'
(RHZ07/MAI.10322 Rhein-Zeitung, 11.05.2007; In Brasilien ist die Religion ein Markt)

29 Probing the actual interpretations of speakers regarding AAN-phrases including evaluatives seems a worthwhile endeavor beyond this chapter's scope. There, possible prosodic differences, such as stress on the second adjective, might prove insightful with respect to the availability of the 'evaluation-qua-second-adjective-reading'.

The scope dependency of evaluatives is a specific feature of this class, which none of the other notional categories exhibit. In the above example, an interpretation where the church is evaluated as beautiful as a consequence of its being large is not promoted, if at all conceivable. Under the assumption that *groß* 'big' denotes a property which tends to be evaluated positively and with said semantic difference between the two orders in mind, their ratio must be concluded to also be based on scopal configurations rather than on a norm alone. Consider the examples in (23):

(23) a. *ein hübsches eckiges Auto* 'a nice boxy car'
b. *ein schöner schwarzer Hund* 'a beautiful black dog'
c. *ein schnelles eckiges Auto* 'a fast boxy car'
d. *ein großer schwarzer Hund* 'a big black dog'
e. ?/**ein eckiges hübsches/schnelles Auto*
f. ?/**ein schwarzer großer/schöner Hund*

In (23a)–(23b), *little car* and *black dog* are preceded by the evaluative modifiers *nice* and *beautiful*, respectively. In one of the readings for both examples, the speaker describes the car as *nice* qua its shape (*boxy*) and the dog as *beautiful* qua its color. This qua-second-property reading is not accessible for the examples in (23c/d). The speed of a car is not conventionally thought of in terms of the vehicle's shape and the size of a dog not in terms of its color. However, reversing the order of the two adjectives is again clearly degraded as shown in (23e/f) for all four examples. Thus, we are faced with the situation that evaluatives in their canonical or typical position in AAN-phrases allow for two interpretations, causal modification qua second quality adjective and non-causal modification, while only the latter one constitutes an AOR phenomenon in the narrow sense.

To sum up, an analysis of AORs as preferences based on norms has an empirical foundation but can be reasonably applied only to semantically identical orders. The question remains of how the occurrence of varying, semantically identical orders of adjective combinations can be accounted for and, furthermore, how to explain the noticeable markedness perceived with reversed orders in out-of-the-blue contexts, given their linguistic acceptability. These issues will be addressed in the following.

4.2 Reversed adjective orders and discourse linking

Under the assumption that order variations among quality adjectives that do not have a semantic rationale are rooted in norm-based preferences, how can

we account for the occurrence of reversed orders? Are they accidental? A factor known to have an impact on adjective orders concerns the domain of information structure. As mentioned above (see fn. 10), focal stress produces adjective orders that are otherwise highly marked. Thus, with focal stress on the first adjective used to indicate a contrastive reading, an otherwise non-preferred order like color preceding dimension adjectives, as in *the YELLOW small flower*, is rendered acceptable. Another information-structural factor that can determine adjective orders is related to discourse linking. Consider the following examples for the untypical orders *weiß* 'white' > > *rund* 'round' as well as *schwarz* 'black' > > *klein* 'small' from our corpus search.

(24) *weiß* > > *rund*
Statt scharfer Augen aus dem All genügt Forschern aber seit dem 19. Jahrhundert auch eine weiße runde Platte von etwa 10 Zentimetern Durchmesser, die an Schnüren ins Wasser gelassen wird.
'Instead of sharp eyes from space, a white round plate of about 10 centimeters in diameter, which is placed into the water on cords, is sufficient for the researchers.'
(SPK/J10.00551 spektrumdirekt, 28.07.2010; Dem Grün nicht grün)

(25) *schwarz* > > *klein*
Nun bekam ich eine hellgrüne Schlafcouch, einen kleinen dreieckigen Tisch mit einer hellgrünen Platte, zwei mittelgrüne Cocktailsessel mit Holzlehnen und schwarzen Ornamenten im Bezug, eine dreiarmige Stehlampe mit drei verschiedenen Farben (grün, orange, gelb), einen schwarzen kleinen Teppich (in den Ecken mit gelb-, grün- und orange-farbigen Ornamenten in den Farben der Lampe), [...]
'Now I got a light green sofa bed, a small triangular table with a light green plate, two medium green lounge chair with wooden arms and black ornaments on the cover, a three-arm floor lamp with three different colors (green, orange, yellow), a black small rug (with yellow, green and orange-colored ornaments in the colors of the lamp in the corners), [...]'
(BRZ07/JUL.15454 Braunschweiger Zeitung, 16.07.2007)

The example in (24) centers around the visibility of an object – a plate used in water to estimate plankton density. The notion of visibility is referred to in the discourse preceding the DP in question through the prepositional phrase *statt scharfer Augen* 'instead of sharp eyes'. The example in (25) contains a list of furniture items, enumerating their color attributes. Considered from a discourse-theoretical perspective, the notions of visibility and color (of furniture items) in

the above examples correspond to so-called discourse topics (see, for example, Asher 2004; Reinhart 1980), i.e. to the subject matter a text is about. Coherence in discourse is achieved through the successful establishment of discourse relations such as *Elaboration*(α,β) or *Narration*(α,β), which progress the content of a text in specific ways. A common definition for *Elaboration* holds that two discourse units *α* and *β* describe the same state of affairs but the latter does so in a more specific way (see, among others, Hobbs 1979: 73–74) and the second argument *β* is a mereological part of the first (see Asher and Lascarides 2003: 161). We assume this to be instantiated in the example in (24): Visibility (*α*) is defined as a function of contrast and the measure of the distance at which an object can be discriminated. Thus, it is determined by factors like size, sharpness as well as contrast with the background, and whiteness (*β*) represents a part of this notional set as it coincides with high contrast.[30] Arguably, the color adjective is thus fronted in order to link the corresponding DP *weiße runde Platte* 'white round plate' to the discourse and produce an increase in coherence. An analogous reasoning can be applied to the example in (25), in which the respective DP *schwarzen kleinen Teppich* 'black little rug' is part of a listing of items stating object properties from the taxonomic field of color. Note that in the DP *einen kleinen dreieckigen Tisch* 'a small triangular table' in the example, the object's color is not denoted and the two adjectives follow the expected order dimension > > shape.

We conclude that factors rooted in discourse linking can be used to account for deviations from preferred adjective orders. Crucially, in the cases in question, conformity to discourse conditions goes hand in hand with linguistic unmarkedness as otherwise non-preferred orders are rendered unproblematic in matching contexts. The question remains of how to explain the markedness perceived with reversed orders in out-of-the-blue contexts. We will address this issue in the following section.

4.3 A pragmatic account of the markedness of reversed orders

The following examples again illustrate the noticeable markedness of reversed orders in out-of-the-blue and wide focus contexts, respectively, that form the basis of many of the primary introspective entries to the field (see also example 1 above).

30 Note that while the shape of an object (as encoded in the adjective *rund* 'round' here) certainly has to do with visual perception, it is not a matter of visibility – the contrast needed for estimating plankton density is provided by the plate's color, not its shape.

(26) a. ?/*Es war einmal ein blauer kleiner Elefant.
'Once upon a time there was a blue small elephant.'
b. *Was ist passiert?*
?*Max hat einen gelben runden Tisch gekauft.*
'What happened?' 'Max bought a yellow round table.'

Observe that the two examples appear to differ in their degree of acceptability, with the example in (26b) less marked than the one in (26a). The contrast may be due to the fact that dimension adjectives like *small* and color adjectives (*blue*), as used in (26a), represent more distant classes on the hierarchy of notional classes (as formulated, for example, by Scott 2002) than the neighboring classes shape (*round*) and color (*yellow*), as in (26b). Such an approach has been adopted by Adam and Schecker (2011), who show on the basis of an ERP-study that adjectives from adjacent classes can be reversed more easily than adjectives from more distant classes. Based on their interpretation of the elicited event-related potential (a LAN), the authors opt for an implementation of the observed effects in the realm of syntax. This, however, is not a necessary conclusion as the LAN is an ERP component that has been reported to be sensitive also to working memory load (see, for example, King and Kutas 1995) and, recently, also to infringements rooted in the domain of pragmatics (see Kulakova et al. 2014).

We believe that a pragmatic account of the markedness perceived with reversed orders like those in (26) offers a theoretically coherent way of implementing AORs as preferences, that is, regularities based on norms.[31] Knowledge about what is the norm (and what is normal) derives from generalizations across typical instances of the entities in the world (see d'Avis 2013), that is, in our case, across occurrences of multiple prenominal adjectives. It represents our understanding about defaults (see Leslie 2008). An important property of this kind of knowledge is, as stressed by d'Avis (2013), that it has to allow exceptions. Depending on the strength of the norm, exceptions can be unmarked or marked. For example, a deviation from the norm that ducks lay eggs is unmarked as more than half of the population of ducks does not lay eggs (cf., among others, Khemlani et al. 2007). A

31 A reviewer asked why we do not treat adjective order preferences as based on order rules on a par with rules that determine sentential word order and information-structural operations like scrambling. Observe, however, the considerable difference in acceptability between deviations from the canonical word order in, e.g. German subordinate clauses (cf. **Tim fragt sich, ob Max mag Kuchen* 'Tim wonders whether Max likes cake') and the adjective order reversals in (26). We stay agnostic w.r.t. the question whether scrambling should indeed be analyzed using the notion of norm as well as the M-principle, as we suggest for adjective order reversals.

marked exception would be a duck that cannot swim as an instance of a deviation from the default that ducks can swim. In the linguistic realm, the markedness of an expression has been linked to its deviation from a more conventionalized form. For example, according to Levinson (2000: 137), an expression is marked if it is less lexicalized. Importantly, the use of a marked expression implicates a non-stereotypical meaning.

This is the gist of Levinson's Manner-principle, which states that a non-stereotypical situation is indicated "by using marked expressions that contrast with those you would use to describe the corresponding normal [...] situation" (Levinson 2000: 136). The M-principle can be used to explain, for example, interpretational differences between NN-compounds, such as *matchbox*, embodying the default interpretation as a specific kind of box, and corresponding phrasal expressions, like *box for matches*, which denotes some non-prototypical box used for matches (see Levinson 2000: 147). In a similar way, we can use the M-principle to account for occurrences of non-standard adjective orders that, as discussed above, are rooted in information-structural domains as illustrated in (24) and (25). In these cases, the choice of a non-stereotypical order creates the implicature of a specific contextual condition to be considered on the hearer's side. Crucially, a manner-based implicature of this sort fails to emerge in out-of-the-blue contexts; see the examples in (26) above. In these cases, an untypical expression, that is, a non-preferred adjective order is used in a situation that does not require a special marking, thus producing the relatively strong markedness effect we perceive in these cases.

The notion of norm as used to account for the markedness of untypical adjective orders implies that default adjective orders are the result of generalizations over occurrences. It could be tempting, then, to ground the derivational mechanism that produces an order for multiple prenominal adjectives on token-based analogy. Such an approach builds on patterns stored in the mental lexicon based on exemplar frequency, with the choice between two alternatives dependent on the frequency value of a pattern (see, for example, Baayen et al. 2010; Schlücker and Plag 2011). The degree of conventionalization of a pattern is essentially a statistical matter in such an approach, in which rule-based, categorial factors are regarded as irrelevant. An alternative approach, which we favor, is to implement categorial (semantic-conceptual) properties of adjectives as an informational source which is exploited in parsing multiple adjective occurrences. A parsing strategy of this kind predicts conceptually more specific adjectives to appear after less specific ones. Thus, the semantic range of options for the positions to follow in a sequence of adjectives is limited. In this vein, we follow Eichinger (1991), who associates adjective order preferences with memory capacities and a strategy of the cognitive system to reduce complexity.

5 Conclusions

Our analysis has shown that – restricting the object of investigation to AORs in the narrow sense based on truth-conditional consistency – only the distinction between object- and kind-modification can be described as a hard constraint, which presumably is located in syntax proper. By default, this bipartition accounts for the order quality > > relational adjective. With respect to object-modifying quality adjectives, we have argued on both a theoretical and empirical basis that the myriad of seemingly (un)grammatical examples derive from norm-based preferences rather than grammatical rules. The corpus study has provided clear-cut evidence that the relative > > absolute hierarchy in German (with its prenominal modifying template) is a statistically significant predictor for said preferences, while neither the temporariness of adjectival concepts nor morphophonological weight are determining factors (the latter, however, possibly due to the not well suited design of the study). In particular, temporariness has been shown to be completely absorbed in more general hierarchy preferences. Although dismissed as core syntactic phenomena, pace the cartographic program, notional classes most likely do play a role in the establishment of norm-based preferences.

Finally, we have argued that several cases of untypical adjective orders can be explained by discourse-linking phenomena and that the markedness of AAN-phrases in out-of-the-blue and wide focus contexts is based on knowledge of norms and the non-applicability of pragmatic repair principles. An open question is if order preferences we observe correlate to conceptual scales as part of our non-linguistic knowledge on a higher cognitive level. Such an assumption could imply cognitive orders to be hardwired into our conceptual system to a significant extent. An alternative hypothesis could be to reverse causality and view cognitive orders as mirroring configurations found in a language and, thus, as largely dependent on the linguistic input and its frequency. Supposed cognitive scales would then be language-specific in nature. This is a chicken-or-the-egg dilemma: While analogy-based accounts see input frequency as the origin of cognitive scales, the opposite view recognizes a categorial rule system as the force behind the formation of linguistic units – and their frequencies as an effect. Note that what is essentially at stake here is the question if our non-linguistic knowledge is dependent on (a specific) language, i.e. the linguistic relativity hypothesis. While this subject is beyond the current paper's focus, the numerous ordering regularities to be found cross-linguistically let us opt for a non-relativistic standpoint, where order preferences are a reflection of conceptual configurations hardwired into the cognitive system, though certainly not as part of universal grammar.

References

Adam, Séverine & Michael Schecker. 2011. Position und Funktion: Kognitive Aspekte der Abfolge attributiver Adjektive. In Günter Schmale (ed.), *Das Adjektiv im heutigen Deutsch. Syntax, Semantik, Pragmatik* (= Eurogermanistik 29), 157–172. Tübingen: Stauffenburg.

Arsenijević, Boban, Gemma Boleda, Berit Gehrke & Louise McNally. 2014. Ethnic adjectives are proper adjectives. In Rebekah B. Baglini, Timothy Grinsell, Jonathan Keane & Julia Thomas (eds.), *Proceedings of CLS 46*, 17–30. Chicago: The Chicago Linguistic Society.

Asher, Nicholas. 2004. Discourse topic. *Theoretical Linguistics* 30(2–3). 163–201.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.

Bache, Carl & Niels Davidsen-Nielsen. 1997. *Mastering English*. Berlin: Mouton de Gruyter.

Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.

Bouchard, Denis. 2002. *Adjectives, number and interfaces: Why languages vary*. Amsterdam: Elsevier.

Bouchard, Denis. 2011. Review: Syntax: Cinque (2010). *Linguist List book review*https://linguistlist.org/issues/22/22-2513.html (June 20, 2016).

Carlson, Greg N. 1980. *Reference to kinds in English*. New York/London: Garland Publishing.

Carlson, Greg N. 2003. Interpretive asymmetries in major phrases. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar: Syntax and semantics*, 301–314. Amsterdam: John Benjamins.

Champollion, Lucas. 2006. A game-theoretic account of adjective ordering restrictions. University of Pennsylvania manuscript, 1–16.http://ling.upenn.edu/~champoll/adjective-ordering.pdf (July 1, 2016).

Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raffaella Zanuttini (eds.), *Paths towards universal grammar. Studies in honor of Richard S. Kayne*, 85–110. Washington D.C.: Georgetown University Press.

Cinque, Guglielmo. 2010. *The syntax of adjectives. A comparative study*. Cambridge, MA/London: MIT Press.

Cooper, William E. & John R. Ross. 1975. World order. In Robin E. Grossman, L. James Sam & Timothy J. Vance (eds.), *Papers from the Parasession on Functionalism*, 63–111. Chicago: Chicago Linguistic Society.

D'Avis, Franz. 2013. *Normalität und Sprache - Normalvorstellungen und ihre Rolle in bestimmten Konstruktionen des Deutschen*. Mainz, Germany: Johannes Gutenberg Universität habilitation treatise.

Dixon, Robert M. W. 1982. *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: De Gruyter.

Eichinger, Ludwig M. 1991. Woran man sich halten kann: Grammatik und Gedächtnis. *Jahrbuch Deutsch als Fremdsprache* 17. 203–220.

Eichinger, Ludwig M. 1992. Ganz natürlich – Aber im Rahmen bleiben. Zur Reihenfolge gestufter Adjektivattribute. *Deutsche Sprache* 4(91). 312–329.

Eroms, Hans-Werner. 2000. *Syntax der deutschen Sprache*. Berlin: de Gruyter.

Fernald, Theodore B. 2000. *Predicates and temporal arguments*. New York/Oxford: Oxford University Press.

Gunkel, Lutz & Gisela Zifonun. 2009. Classifying modifiers in common names. *Word Structure* 2(2). 205–218.

Halliday, Michael A. K. 2014. *Halliday's introduction to Functional Grammar* 4th edn revised by Christian Matthiessen. London: Routledge.

Härtl, Holden. 2016. Normality at the boundary between word-formation and syntax. In Franz d' Avis & Horst Lohnstein (eds.), *Normalität in der Sprache* (= Linguistische Berichte Sonderhefte 22), 71–98. Hamburg: Buske.

Hetzron, Robert. 1978. On the relative order of adjectives. In Hansjakob Seiler (ed.), *Language Universals*, Papers from the Conference held at Gummersbach/ Cologne,Germany, October 3–8, 1976, 165–184. Tübingen: Narr.

Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1). 67–90.

Jäger, Gerhard. 2001. Topic-comment structure and the contrast between stage level and individual level predicates. *Journal of Semantics* 18(2). 83–126.

Kamp, Hans. 1975. Two theories about adjectives. In Edward L. Keenan (ed.), *Formal semantics of natural language*, 123–155. Cambridge: Cambridge University Press.

Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81. 345–381.

Khemlani, Sangeet, Sarah-Jane Leslie, Sam Glucksberg & Paula Rubio-Fernandez. 2007. Do ducks lay eggs? How people interpret generic assertions. In Danielle S. McNamara & J. Greg Trafton (eds.), *Proceedings of the 29th Annual Cognitive Science Society Austin*, 395–400. TX: Cognitive Science Society.

King, Jonathan W. & Marta Kutas. 1995. Who did what to when?: Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7(3). 376–395.

Kotowski, Sven. 2016. *Adjectival modification and order restrictions. The influence of temporariness on prenominal word order* (= Studia Grammatica 80). Berlin: de Gruyter.

Kotowski, Sven & Holden Härtl. 2011. Recursion and the language faculty. On the evolution of the concept in generative grammar. In Said Sahel & Ralf Vogel (eds.), *NLK-Proceedings: 12. Norddeutsches Linguistisches Kolloquium* (NLK2011 Braunschweig), 24–47. Bielefeld: Bielefeld eCollections.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The Generic Book*, 125–175. Chicago: The University of Chicago Press.

Kulakova, Eugenia, Dominik Freunberger & Dietmar Roehm. 2014. Marking the counterfactual: ERP evidence for pragmatic processing of German subjunctives. *Frontiers in Human Neuroscience* 8.http://journal.frontiersin.org/article/10.3389/fnhum.2014.00548 (December 5, 2016).

Laenzlinger, Christopher. 2005. French adjective ordering: Perspectives on DP-internal movement types. *Lingua* 115. 645–689.

Larson, Richard K. 1998. Events and Modification in Nominals. In Devon Strolovitch & Aaron Lawson (eds.), *Proceedings from Semantics and Linguistic Theory (SALT) VIII*, 145–168. Ithaca, NY: Cornell University.

Leslie, Sarah-Jane. 2008. Generics: Cognition and acquisition. *The Philosophical Review* 117(1). 1–49.

Levi, Judith. 1978. *The syntax and semantics of complex nominals*. New York: Academic Press.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Maienborn, Claudia. 2003. *Die logische Form von Kopula-Sätzen* (= Studia Grammatica 56). Berlin: Akademie Verlag.

McNally, Louise & Gemma Boleda. 2004. Relational adjectives as properties of kinds. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics* Vol. *5*, 179–196.

Published only electronically: http://www.cssp.cnrs.fr/eiss5/mcnally-boleda/mcnally-boleda-eiss5.pdf (accessed 05 January 2019)

Motsch, Wolfgang. 2002. Die Wortart 'Adjektiv'. In D. Alan Cruse, Franz Hundsnurscher, Michael Job & Peter Rolf Lutzeier (eds.), *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen* (1st sub-vol.), 598–604. Berlin: de Gruyter.

Motsch, Wolfgang. 2004. *Deutsche Wortbildung in Grundzügen*. 2nd ed. Berlin: de Gruyter.

Partee, Barbara H. 1995. Lexical semantics and compositionality. In Daniel Osherson, Lila Gleitman & Mark Liberman, (eds.), *Invitation to cognitive science. Part I: Language*, 311–360. Cambridge, MA: MIT Press.

Payne, John & Rodney Huddleston. 2002. Nouns and noun phrases. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 323–523. Cambridge: Cambridge University Press.

Posner, Roland. 1980. Ikonismus in der Syntax. Zur natürlichen Stellung der Attribute. *Zeitschrift für Semiotik* 2. 57–82.

R. Harald, Baayen, Victor Kuperman & Raymond Bertram. 2010. Frequency effects in compound processing. In Sergio Scalise & Irene Vogel (eds.), *Cross-Disciplinary Issues in Compounding*, 257–270. Amsterdam: John Benjamins.

Ramaglia, Francesca. 2011. *Adjectives at the syntax-semantics interface*. München: Lincom Europa.

Reinhart, Tanja. 1980. Conditions for text coherence. *Poetics Today* 1(4). 161–180.

Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.

Rijkhoff, Jan. 2010. Functional categories in the noun phrase: On jacks-of-all-trades and one-trick-ponies in Danish, Dutch and German. *Deutsche Sprache* 2. 97–124.

Rotstein, Carmen & Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12(3). 259–288.

Schlücker, Barbara. 2014. *Grammatik im Lexikon. Adjektiv-Nomen-Verbindungen im Deutschen und Niederländischen* (= Linguistische Arbeiten 553). Berlin & Boston: de Gruyter.

Schlücker, Barbara & Ingo Plag. 2011. Compound or phrase? Analogy in naming. *Lingua* 121. 539–551.

Schurz, Gerhard. 2001. What is 'normal'? An evolution-theoretic foundation for normic laws and their relation to statistical normality. *Philosophy of Science* 68(4). 476–497.

Scontras, Gregory, Judith Degen & Noah D. Goodman. 2017. Subjectivity predicts adjective ordering preferences. *Open Mind* 1. 53–65.

Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In Guglielmo Cinque (ed.), *Functional structure in the DP and IP: The cartography of syntactic structures* (vol. I), 91–120. Oxford: Oxford University Press.

Seiler, Hansjakob. 1978. Determination: A functional dimension for interlanguage comparison. In Hansjakob Seiler (ed.), *Language Universals*, Papers from the Conference held at Gummersbach/ Cologne,Germany, October 3–8, 1976, 301–328. Tübingen: Narr.

Solt, Stephanie. 2016. Ordering subjectivity and the absolute/relative distinction. In Nadine Bade, Polina Berezovskaya & Anthea Schöller (eds.), *Proceedings of Sinn und Bedeutung* Vol. *20*, 676–693.
Published only electronically: https://semanticsarchive.net/sub2015/SeparateArticles/Solt-SuB20.pdf (accessed 05 January 2019)
Sproat, Richard & Chilin Shih. 1988. Prenominal adjectival ordering in English and Mandarin. In James Blevins & Juli Carter (eds.), *Proceedings of NELS 18*, 465–489. Amherst, MA: GLSA.
Svenonius, Peter. 1994. The Structural location of the attributive adjective. In Eric Duncan, Donka Farkas & Philip Spaelti (eds.), *Proceedings of WCCFL 12*, 429–454. Stanford: CSLI Publications.
Svenonius, Peter. 2008. The position of adjectives and other phrasal modifiers in the decomposition of DP. In Louise McNally & Chris Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*, 16–42. Oxford: Oxford University Press.
Teodorescu, Alexandra. 2006. Adjective ordering restrictions revisited. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 399–407. Somerville, MA: Cascadilla Proceedings Project.
Trost, Igor. 2006. *Das deutsche Adjektiv*. Hamburg: Buske.
Truswell, Robert. 2009. Attributive adjectives and nominal templates. *Linguistic Inquiry* 40. 525–533.
Van der Hulst, Harry. 2010. Re recursion. In Harry van der Hulst (ed.), *Recursion in Human Language*, xv–liii. Berlin & New York: Mouton de Gruyter.
Vendler, Zeno 1968. *Adjectives and nominalizations*. The Hague: Mouton & Co.
Vergnaud, Jean-Roger & Maria Luisa Zubizarreta. 1992. The definite determiner and the inalienable constructions in French and English. *Linguistic Inquiry* 23(4). 595–652.
Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Berlin & New York: Walter de Gruyter.