

Ad-hoc kind formation by similarity¹

Carla Umbach & Britta Stolterfoht

1 Introduction

This paper combines three topics that are not initially thought to be connected, which are (a) the semantics of *demonstratives of manner, quality and degree*, (b) the notion of *in-virtue-of properties*, and (c) the idea of *ad-hoc formation of kinds* by linguistic means.

Demonstratives expressing manners, qualities and degrees occur across languages including, e.g., German *so/solch*, English *such* and Polish *tak* (see König and Umbach 2018). Following Umbach and Gust (2014), they are genuine demonstratives denoting, however, similarity instead of identity – this is why we call them *similarity demonstratives*. For example, the meaning of *so ein Tisch* in (1) is not the table the speaker points at but instead a table similar to the one pointed at.

- (1) a. (speaker points at a table):
 So einen Tisch hat Berta auch.
 'Berta has such a table / a table like this, too.'

For semantic interpretation, Umbach and Gust (2014) define similarity as indistinguishability with respect to a set of features of comparison. However, the features that are suited for similarity comparison are restricted. In (1), for example, the phrase *so ein Tisch* 'such a table' may denote a table with three-legs and a glass top, but not a table formerly owned by Aunt Elsi. Obviously, features of comparison may include the number of legs and the type of the top, but not the former owner.

Restrictions on features permissible in similarity comparison were subject to a series of experimental tests reported in this paper. We tested nominal as well as verbal similarity phrases (*so ein Tisch* / *so tanzen* 'a table like this / 'dance in this way'). It was found that permissible features of comparison correspond to internal modifiers (in the sense of Maienborn 2003 as well as Specht and Stolterfoht 2022), which affect kinds instead of individuals or events, e.g., *auf den Fingern pfeifen* 'whistle on one's fingers'. See also the two ways of modification described by McNally & Boleda (2004) reflecting kind modification vs. modification of individuals. However, there is no explanation to be found in the literature as to why certain properties qualify as features of comparison or internal modifiers, *rsp.*, and others do not.

This question leads to the second topic of this paper, which is the notion of *in-virtue-of properties*. It plays a major role in the research of Prasada and colleagues in cognitive psychology (e.g., Prasada & Dillingham 2006, Prasada et al. 2013) who provide experimental evidence showing that humans represent principled connections between concepts and particular properties, for example the property of 'having four legs' for the concept DOG. Principally connected properties license explanations in terms of in-virtue-of properties, for example, *dogs have four legs in virtue of being*

¹ Intermediate results of the research reported in this paper were presented at workshops and colloquia in Frankfurt, Rijeka, Bochum, Bar Ilan and at the DGfS 2024. We would like to thank the audiences for their helpful discussions and encouragement, in particular Louise McNally, Stephanie Solt, Sebastian Bücking, Sandeep Prasada and Vera Hoffmann-Kolss. We would also like to thank Felix Frühauf and Stella Krüger for their creativity in designing stimuli and conducting exploratory studies, and Tamara Stoyke and Hanna Blach for their help with running the final study and translating the stimuli. Finally, we are grateful to an anonymous reviewer for very helpful comments, and to the editors of this volume, Peter Sutton and Manfred Krifka, for their support and patience. The authors acknowledge funding by the Deutsche Forschungsgemeinschaft, DFG UM 100/1-3 and SFB 833/B8.

dogs. Greenberg (2007), moreover, observes that they license indefinite singular generic sentences – *a dog has four legs*. In-virtue-of properties are relevant in this paper because there seems to be a close, albeit not direct, link to the features permissible in similarity comparison – a table may have a glass top in virtue of being a table, but it cannot, in virtue of being a table, belong to Aunt Elsi.

The third topic of this paper is the idea of *ad-hoc formation* of kinds. The link between features of comparison and in-virtue-of properties is reason to predict that similarity phrases like *solche Tische* 'tables like this' establish ad-hoc subkinds – the kind of tables similar to the one pointed to. It has already been argued by Carlson (1977) and subsequently by Anderson and Morzycki (2015) that phrases headed by *such* refer to kinds. The main criticism of this idea is that it ignores the semantic contribution of *such* obscuring the understanding of how these kinds come about and how this process is restricted.

The topic of ad-hoc formation was launched by Barsalou's (1983) paper on ad-hoc categories which are collections of items (individuals, actions etc.) relevant for achieving specific goals (e.g., *how to escape from the mafia*). Mauri (2017) presents a cross-linguistic survey of strategies of forming ad-hoc categories, including open-ended lists (*water, herbal teas, and the like*) and simulative plurals (as in Turkish reduplication, e.g., *havlú-mavlú* 'towels and the like'). Focusing on kinds, the idea of ad-hoc formation triggers a number of interesting issues in the field of genericity – how to decide, for example, whether open-ended lists denote kinds or just plural individuals? And is there reason to assume that ad-hoc kinds are generally based on similarity?

In section 2 of this paper, we give a brief overview over the analysis of similarity demonstratives, which we take as a basis here. In section 3 experimental evidence will be provided showing that features permissible in similarity comparison are restricted to those allowing for internal modification. In section 4, the notion of in-virtue-of properties is discussed asking whether they correspond to features permissible in similarity comparison, and whether similarity classes denote kinds. In section 5 we will consider ad-hoc kind formation beyond similarity phrases. In the conclusion we take a brief look at the notion of intrinsic properties in philosophy which seems to match surprisingly well with our findings.

2 The semantics of similarity demonstratives

2.1 Similarity instead of identity

Demonstratives expressing manner, quality and degree – we call them *similarity demonstratives* – occur across languages including German *so* and *solch*, English *such*, Polish *tak*, Italian *così*, Hebrew *kaze*, Turkish *öyle/ böyle/ şöyle* and many more (see König & Umbach 2018). They function as modifiers of nominal, verbal, and/or adjectival phrases and can be used deictically as well as anaphorically. The analysis presented in this paper takes German *so* as its starting point. In (2) it is used deictically, expressing a quality in (a), a manner in (b) and a degree in (c). In (3), the corresponding anaphoric uses are shown.²

(2) a. (speaker points at a table)

So einen Tisch hat Berta auch.

'Berta has such a table / a table like this, too.'

²The demonstrative *so* also occurs in equative comparison constructions which may be considered as a cataphoric use with the demonstrative picking up the standard of comparison, (a). If the standard includes a regular demonstrative, the equative in (b) is equivalent to the deictic use of *so* in (2a).

(a) *Berta hat so einen Tisch wie der von Anna*. 'Berta has a table like Anna's.'

(b) (Speaker pointing at a table) *Berta hat so einen Tisch wie dieser (einer ist)*. 'Berta has a table like this one.'

- b. (speaker points to someone dancing)
So tanzt Berta auch. 'Berta is dancing like this, too.'
 - c. (speaker points to a person)
So groß ist Berta auch. 'Berta is this tall, too.'
- (3) a. Anna hat einen Tisch mit drei Beinen. 'Anna has a three-legged table. Berta has such a
So einen Tisch hat Berta auch. table / a table like this, too.'
- b. Anna tanzt mit wilden Sprüngen. 'Anna is dancing with wild spins. Berta is dancing
So tanzt Berta auch. like this, too.'
 - c. Anna ist 1,80m groß. 'Anna is 1.80. Berta is this tall, too.'
- So groß ist Berta auch.

In all of these cases, the manner or quality or degree denoted by the *so*-phrase has to be inferred by similarity – Berta's table is similar (in certain respects) to the table pointed at or mentioned in the preceding sentence. This is why these demonstratives are termed *similarity demonstratives*.³ From the point of view of semantics, they pose the problem of how to reconcile their demonstrative characteristics with their modifying capacity. Following Umbach & Gust (2014), they are directly referential in the sense of Kaplan (1989) pointing to individuals or events. However, while standard demonstratives express identity between the demonstration target and the referent (as is in-built in Kaplan's system), similarity demonstratives express similarity rather than identity. This is why *so ein Tisch / such a table* does not refer to the table the speaker points at but instead to a similar one.⁴

The notion of similarity plays a prominent role in Cognitive Science explaining cognitive skills like perception and classification. Following Tversky (1977), "similarity [...] serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations." (p. 327). But the similarity relation in itself is not unproblematic. Goodman (1972) argued that "... similarity cannot be equated with the possession of common characteristics [...] since every two objects have infinitely many properties in common." (p. 443) (for example, a laptop and an orange both weigh less than 100kg, 101kg, 102kg, etc.) He concludes that we have to say in what respect two things are similar. This is the issue of the features permissible in similarity comparison, which is the core topic of this paper.

2.2 The similarity framework

There are in Cognitive Science basically two ways to spell out the relation of similarity, either distance-based, as in Gärdenfors' (2000) conceptual spaces, or feature-based, as in Tversky's (1977) contrast model. In Umbach & Gust (2014)/(2021), a feature-based approach is presented which, unlike Tversky's model, starts from defining similarity as indistinguishability with respect to given features of comparison and granularity. Graded similarity is built on absolute similarity / indistinguishability by refinement in a Kleinian fashion.⁵ Features of comparison correspond to dimensions defining multidimensional attribute spaces. They are equipped with measure functions – we call them *generalized measure functions* – mapping entities of a domain to points in these spaces. Dimensions and generalized measure functions are two components of what is called a *representation*. The third

³ In König & Umbach (2018) they are called *MQD demonstratives* – demonstratives of manner, quality and degree.

⁴ It was already observed by Nunberg (1993) that demonstratives do not presuppose identity of target and referent. Nunberg focuses on contiguity, as in *This is parked around the corner* (speaker handing a car key to the addressee) which is enforced by context. In contrast, the similarity interpretation of *so* is lexically determined.

⁵ In Klein (1980) the comparative of a gradable adjective is built upon its positive form by refinement of contexts, e.g., *A is smaller than B* iff there is a context in which A counts as small while B does not.

component is a set of *classifiers*, which are predicates on points in attribute spaces fixing granularity by defining a "grid" of indistinguishable points. Slightly simplifying, a representation $\mathcal{F}$ is defined as a quadruple including a domain D , an attribute space F , a generalized measure functions $\mu: D \rightarrow F$ and a set of classifiers P^* , $\mathcal{F}_D = \langle F, \mu, P^*, D \rangle$. All of the parameters are related to the type of the domain, so we add indices when necessary to avoid confusion (for formal definitions and details of the framework see the appendix in Umbach & Gust 2021).

Similarity is defined as a three-place relation combining two entities to be compared and a representation, $SIM(x, y, \mathcal{F}_D)$, such that the entities x and y are similar relative to the representation $\mathcal{F}_D$ if and only if the points in the attribute space they are mapped to are indistinguishable relative to the set of classifiers. Similarity defined in this way is an equivalence relation.⁶

In (4), the semantic interpretation of *so einen Tisch / such a table* is shown. Let us assume, for the sake of the example, that relevant dimensions of comparison are HEIGHT, MATERIAL, LEGS, EXTRAS and VINTAGE, and that tables are "measured" by the function in (5). Now suppose that the table the speaker points at is mapped to $\langle 55cm, metal, 4, \{\}, + \rangle$ and the set of classifiers is such that points within a range of $HEIGHT:40-60; MATERIAL:\{metal, plastics\}; LEGS:2-4; EXTRAS:\{\}, VINTAGE: \{+, -\}$ cannot be distinguished. Then (2a) is true iff Berta's table is mapped to a point within this range.

$$(4) \quad [[\text{so ein Tisch} / \text{such a table}]] = \lambda x. \text{table}(x) \ \& \ SIM(x, t, \mathcal{F}_{\text{table}})$$

$$(5) \quad \mu_{\text{table}} : D_{\text{table}} \rightarrow \text{HEIGHT} \times \text{MATERIAL} \times \text{LEGS} \times \text{EXTRAS} \times \text{VINTAGE}$$

where $\mu_{\text{table}}(x) = \langle \mu_{\text{height}}(x), \mu_{\text{material}}(x), \mu_{\text{legs}}(x), \mu_{\text{extras}}(x), \mu_{\text{vintage}}(x) \rangle$

and

HEIGHT:	$\mathbb{R}^+$
MATERIAL:	$\{\text{WOOD, METAL, PLASTICS, ...}\}$
LEGS:	$\{1 \dots 10\}$
EXTRAS:	$\wp\{\text{EXTENDABLE, HEIGHT-ADJUSTABLE, COATING, ...}\}$
VINTAGE:	$\{+, -\}$

One might wonder about the compatibility of this system with standard referential semantics which is essential for natural language interpretation. It should be noted, however, that this system can be seen as a generalized version of degree semantics. In degree semantics, measure functions are one-dimensional, mapping entities to values on a metrical scale. Generalized measure functions include multiple dimensions mapping entities to values in each dimension, that is, points in multi-dimensional spaces. Unlike degree measure functions, generalized measure functions allow for arbitrary scale levels: scales may be metrical (like HEIGHT and LEGS in (5)), but also ordinal (EXTRAS) or nominal (MATERIAL) or even binary (VINTAGE).

2.3 The interpretation of similarity phrases

The similarity relation defined above serves in interpreting similarity phrases, that is, phrases in which similarity demonstratives occur as modifiers, as in (2) and (3). Similarity adjectives (*ähnlich / similar, gleich / same*) are excluded, because even though their meaning is also based on the relation of similarity they are subject to additional constraints.⁷

In the case of nominal and verbal phrases, similarity comparison is multi-dimensional, as demonstrated in the example in (5), and features of comparison are implicit (as long as there is no

⁶ Counter-arguments (going back to Tversky 1977) against defining similarity as an equivalence relation cannot in general be maintained, see Umbach & Gust (2014).

⁷ The difference in usage between *so* 'such', *ähnlich* 'similar' and *gleich* 'same' are investigated in Umbach & Gust (2021). A related investigation addressing English *like, similar and resemble* is presented in Fishman (2020).

linguistic specification like *in terms of the number of legs* etc.). In adjectival phrases like *so groß* 'tall like this', there is mostly just one metrical dimension,⁸ and the feature of comparison is given by the meaning of the adjective. This is why adjectival similarity phrases are irrelevant to the issue of permissible features of similarity comparison, and are excluded in this paper. Since the empirical studies in section 3 are conducted in German, we focus on German similarity phrases: nominal phrases of the form *(ein) so/solch (ein) N* and verbal phrases of the form *so V*.

There are two options for an interpretation of adnominal *so/such* in nominal (generalized) quantifiers: Either *so/such* are considered as modifiers of the indefinite determiner, or they are considered as modifiers of the nominal (and may have been moved into the prenominal position). The first option yields the interpretation in (a) and the second the one in (b). The resulting quantifiers are identical.

- (6) a. *so/solch ein Tisch* 'such a table'
 $[[so\ ein]] = \lambda P. \lambda Q. \exists x. SIM(x, t, \mathcal{F}_P) \ \& \ P(x) \ \& \ Q(x)$ ⁹
 $[[so\ ein]] \ ([[Tisch]]) = \lambda Q. \exists x. SIM(x, t, \mathcal{F}_{table}) \ \& \ table(x) \ \& \ Q(x)$
- b. *ein solcher Tisch* 'a such table'
 $[[solch\ Tisch]] = \lambda x. SIM(x, t, \mathcal{F}_{table}) \ \& \ table(x)$
 $[[ein]] \ ([[solch\ Tisch]]) = \lambda Q. \exists x. SIM(x, t, \mathcal{F}_{table}) \ \& \ table(x) \ \& \ Q(x)$

In this paper we ignore quantifiers in favor of a plain predicate interpretation (of type $\langle et \rangle$). Nominal similarity phrases are treated like bare nominals, as in (7a), which is all that is needed for the purposes of this paper. The line in (b) contains a bit of 'syntactic sugar' to emphasize the similarity set or class resulting from the similarity phrase and is equivalent to that in (a).

- (7) *solche Tische* 'such tables'
- a. $[[solch\ Tisch]] = \lambda x. SIM(x, t, \mathcal{F}_{table}) \ \& \ table(x)$
b. $= \lambda x. x \in \{y \mid SIM(y, t, \mathcal{F}_{table}) \ \& \ table(y)\}$

For verbal similarity phrases the interpretation is analogous, as shown in (8).

- (8) *so tanzen* 'dance in this way'
- a. $[[so\ tanzen]] = \lambda e. SIM(e, t, \mathcal{F}_{dance}) \ \& \ dance(e)$
b. $= \lambda e. e \in \{e' \mid SIM(e', t, \mathcal{F}_{dance}) \ \& \ dance(e')\}$

Two remarks on terminology: The examples in (6)–(8) display the deictic use of the demonstrative *so* in which it refers to an item similar to the target of the demonstration gesture (*such a car* is interpreted as a car similar to the target of the demonstration gesture). The target is represented by a contextual parameter t and the features of similarity comparison are implicit. In the anaphoric use of the demonstrative, there is an antecedent instead of the target of the demonstration gesture. If the antecedent expression contains a modifier, the modifying property is preferably understood as the property relevant in comparison (if the antecedent is *car with a gas drive*, then *such a car* is preferably interpreted as a car with a gas drive). Therefore, it looks as though the demonstrative is anaphorically related to the modifier, while in fact it refers to the nominal antecedent (*car with a gas drive*). So the idea that the demonstrative refers to a property is misleading. We nevertheless speak of *so* "picking up a property" for ease of presentation.

⁸ See, however, Sassoon's (2013) analysis of multidimensional adjectives such as *healthy* and *sick*.

⁹ Please note that the index of the representation $\mathcal{F}_P$ starts out as a variable and is fixed when combined with the meaning of the noun.

The other remark concerns the notion of a feature. Technically, features of comparison are not properties, but rather attributes, or dimensions, of a multi-dimensional representation, see the example in (5). A feature of comparison in the context of cars could be COLOR with values {red, green, blue, ...}, or DRIVE-TYPE with values {petrol, gas, electric, ...}. A feature requires a value to turn into a property, as in COLOR:red or DRIVE-TYPE:gas. Features encoding a full blown property are borderline cases; their values are just plus/minus, see VINTAGE in the example in (5).

So again, strictly speaking, the modifier *with a gas drive* is not a feature but a value. However, when speaking of features of similarity comparison, there are two readings: On the strict reading, one looks at genuine features asking whether they have the same value, while on the sloppy reading one looks at properties asking whether they are the same. The two readings do in fact not make a difference except for the definition of *k-features* in section 4.2, which requires the strict reading.

3 Constraints on features of comparison

The features permissible in similarity comparison are, first of all, restricted by the lexical item the similarity demonstrative combines with. In combination with the noun *table*, as in (4), features may be, for example, SIZE, MATERIAL, NUMBER OF LEGS etc. but not NUMBER OF WHEELS or HORSEPOWER. Likewise, the verb *dance* may allow for comparison with respect to POSTURE, EFFORT, and BALANCE but not SOUND VOLUME or VOCAL REGISTER. These restrictions are hardly surprising. It was clear from the outset, however, that there are constraints that cannot be explained by the lexical meaning of the head noun or verb. In Umbach & Gust (2014) we discussed the examples in (9). While A's reply to B in (9a) and (b) appear unmarked, the reply in (9c) is less felicitous. This is obviously due to the fact that *gas drive* refers to the type of engine of cars and being dented is a common appearance of cars, whereas being decorated with a sticker is not in any way specific for cars. At the same time, the use of *so* in (10 B) - intended to refer to the parking ticket - is infelicitous even though parking tickets can only be found with cars.

- (9) A: (pointing to a car in the street):
 So ein Auto ist Anna's Auto auch.
 'Anna's car is one like this.'
 B: In welcher Hinsicht?
 'In which respect?'
 a. A': Anna's Auto hat auch einen Gasantrieb.
 'Anna's car also has a gas drive.'
 b. A': Anna's Auto ist auch vollkommen verbeult.
 'Anna's car is also heavily dented.'
 c. A': ??? Anna's Auto hat auch einen 'I like Berlin' Aufkleber.
 'Anna's car also has an 'I like Berlin' sticker.'
- (10) A: Guck mal, das Auto da drüben hat einen Strafzettel.
 'Look, the car over there has a parking ticket.'
 B: Auf der anderen Straßenseite steht auch so ein Auto.
 'Such a car is on the other side of the street, too.'
 A': Was meinst du?
 'What do you mean?'
 B': # Na, auch eins mit einem Strafzettel.
 'One with a parking ticket, too.'

What is shown in (9c) and (10) is reminiscent of the examples of the infelicitous use of *such* discussed in Carlson (1980), e.g. *people in the next room ... ?? Such people ...*, where the attribute *in the next room* is said to be unfit for selecting a subkind. We propose an explanation why this is the case.

3.1 Pretests

We conducted a series of exploratory pretests considering diverse types of modifiers of nominal as well as verbal items. Since the intended features of comparison can hardly be tested when the demonstrative is used deictically, we switched to anaphoric uses. Test items consisted of a sequence of two sentences such that the first introduces an individual or event with a modifying property. The second sentence contains a similarity demonstrative *so* which is supposed to pick up the modifying property. A continuation by *nämlich...* 'that is' is added to the end of the second sentence in order to make sure that participants interpret *so* in the intended way.¹⁰

In (11) and (12), modification is adnominal. In (11a) the property of Anna's bike presented by the modifying PP is readily picked up by *so*, leading to the interpretation that Berta's bike is also one with gears. In (11b) the property of Anna's bike presented by an attributive modifier is picked up just as readily, leading to the interpretation that Berta's bike is an electric one. In (11c) there are two separate modifiers and in (11d) there is no modifier but instead a more specific noun. Still, the intended characteristics of Anna's bike is readily picked up by the similarity demonstrative in the subsequent sentence.

- (11) a. Anna hat ein Rad mit Gangschaltung. Berta hat auch so ein Rad (nämlich mit Gangschaltung).
 b. Anna hat ein elektrisches Rad. Berta hat auch so ein Rad (nämlich ein elektrisches).
 c. Anna hat ein elektrisches Rad mit Gangschaltung. Berta hat auch so ein Rad (nämlich elektrisch mit Gangschaltung).
 d. Anna hat ein Mountainbike. Berta hat auch so ein Rad (nämlich ein Mountainbike).

'Anna has a bike with gears / an electric bike / an electric bike with gears / a mountain bike. Berta has a bike like that, too (that is, one with gears / an electric one / an electric one with gears / a mountain bike).'

In (12) judgments are more subtle. Can (a) be understood such that Berta has a Greek bike? The problem speakers report with this example is that they don't have a clear picture of Greek bikes. This is different in (b), since (in Germany) Dutch bikes are a well-established kind (heavy, durable, upright sitting position ...). In (c) Anna's bike is said to be a new one, in the sense of *recently purchased*. But an interpretation such that Berta's bike is also new is consistently rejected. Although Berta's bike may share some other property with Anna's bike, being new seems inaccessible for the similarity demonstrative. This observation is confirmed in (d) which can be interpreted such that Berta has a mountain bike, but not such that she has a recently purchased mountain bike.

- (12) a. ?? Anna hat ein griechisches Rad. Berta hat auch so ein Rad (nämlich ein griechisches).
 b. Anna hat ein holländisches Rad. Berta hat auch so ein Rad (nämlich ein holländisches).
 c. # Anna hat ein neues Rad. Berta hat auch so ein Rad (nämlich ein neues).
 d. # Anna hat ein neues Mountainbike. Berta hat auch so ein Rad (nämlich ein neues Mountainbike).

'Anna has Greek / Dutch / new bike / new mountain bike. Berta has a bike like that, too (that is, a Greek / Dutch / new bike / new mountain bike).'

¹⁰ Without *nämlich*-continuation the sentence in (11a), for example, might be understood such that Berta has a bike like Anna's, though maybe without gears.

If, however, *new* is interpreted in the sense of *newly developed* instead of *recently purchased*, it can be picked up by the demonstrative. In (13), Anna bought an exemplar of a novel version of iPhones. This time, the second sentence is preferably interpreted such that Berta has the same novel version.

- (13) Anna geht immer mit der Zeit. Jetzt hat sie sogar ein neues iPhone. Berta hat auch so ein iPhone (nämlich die neueste Version).
'Anna is always up to date. She even has a new iPhone. Berta has such an iPhone, too (that is, the latest version).

Adverbial cases are shown in (14) and (15). The manner of preparing poultry specified in the antecedent sentences is readily picked up by the similarity demonstrative regardless of whether it is expressed by a locative modifier, as in (14a), or by a manner adverbial, as in (b), or by a separate lexeme specifying a manner of preparing food, as in (c). In contrast, the locative modifier in (15a) and the subjective adverbial in (15b) are inaccessible.¹¹

- (14) a. Anna hat das Huhn im Wok zubereitet. Berta hat die Ente auch so zubereitet (nämlich im Wok).
b. Anna hat das Huhn fettarm zubereitet. Berta hat die Ente auch so zubereitet (nämlich fettarm).
c. Anna hat das Huhn gebraten. Berta hat die Ente auch so zubereitet (nämlich gebraten).
'Anna prepared the chicken in the wok / low-fat style / in the frying pan. Berta prepared the duck like this, too (that is, in the wok / low-fat style / in the frying pan).'
- (15) a. # Anna hat das Huhn im Garten zubereitet. Berta hat die Ente auch so zubereitet (nämlich im Garten).
b. # Anna hat das Huhn lecker zubereitet. Berta hat die Ente auch so zubereitet (nämlich lecker).
'Anna prepared the chicken in the garden / reluctantly. Berta prepared the duck like this, too (that is, in the garden / tastily).'

In cases in which the modifier cannot be picked up by *so*, pretests were supplemented by interviews. Participants were asked to think of "repair options", that is, ways how the world would have to be to make access by *so* possible. In (16), for example, the possessive PP *von Tante Anna* 'belonging to Aunt Anna' was rejected by the participants. As a repair it was suggested that if *Tante Anna* were a fashion label, accessibility by *so* would be possible. Analogously, for (17), participants proposed a world where only two companies produce laptops, and one of them is called *Paul*. As in the previous example, the relation of the dress and the laptop, *rsp.*, to another individual – Aunt Anna and Paul – was re-interpreted as a property of the nominal concept – the fashion label *Tante Anna* and the laptop producer *Paul*.

- (16) acceptable with repair: *Tante Anna* refers to a fashion label
Kurz vor dem Abschlussball hat Lisa auf dem Dachboden ein Kleid von Tante Anna gefunden. Ihre beste Freundin hat auch so ein Kleid gefunden, ...nämlich eins von Tante Anna.
'Shortly before the prom, Lisa found a dress of Aunt Anna in the attic. Her best friend also found a dress like this, ... that is, one of Aunt Anna.'

¹¹ In (15b) there is the option of repeating the adverb in combination with *so*: Berta hat die Ente auch so lecker zubereitet 'Berta also prepared the duck so tastily.' In this case, *so* would pick up the degree of tastiness expressed in the first sentence (which is beyond the norm, due to the positive form of the adverb).

- (17) acceptable with repair: *Paul* refers to a laptop manufacturer

Lisa hat sich bei Ebay einen Laptop, der von Paul angeboten wurde, gekauft. Ihr Freund würde sich nie so einen Laptop anschaffen, ...nämlich einen Laptop, der von Paul angeboten wurde.
'Lisa bought a laptop on eBay that was being sold by Paul. Her boyfriend would never buy such a laptop, ... that is, a laptop that was being sold by Paul.'

In the example in (18) *so* is used adverbially. Participants rejected access to the temporal PP *in zwei Stunden* 'in two hours'. However, if the PP is interpreted as a category of decoration with fixed results – decorating a shop in two hours means putting fairy lights in the shop windows, decorating it in three hour means adding garlands on the walls, etc. – access would be unproblematic.

- (18) acceptable with repair: *in two hours* refers to a decoration category

Andrea hat das vordere Schaufenster in zwei Stunden dekoriert. Sven hat den Ladenbereich auch so dekoriert, ...nämlich in zwei Stunden.
'Andrea decorated the front window in two hours. Sven also decorated the store in this way, ... that is, in two hours.'

Finally, in (19) the intended antecedent modifier is a subjective predicate, which is strongly rejected. As a repair option, participants suggested the possibility of expressing the rating of hotels by adjectives instead of numbers of stars – *excellent* = 5 stars, *beautiful* = 4 stars, *nice* = 3 stars, etc. Thus the subjective predicate is re-interpreted as a degree on an objective metrical scale.

- (19) acceptable with repair: *nice* denotes a particular number of stars

Anna zieht sich zum Entspannen gerne in ein schönes Hotel zurück. Ihre Freundin Berta verbringt ihre Ferien auch gerne in so einem Hotel, ...nämlich in einem schönen.
'Anna likes to retreat to a nice hotel to relax. Her friend Berta also likes to spend her vacations in such a hotel, that is, a nice one.'

The repairs are evidence that modifiers related to other individuals turn into accessible ones if they are re-interpreted as subkinds of the kind denoted by the noun or verb. Dresses produced by a particular fashion label are a subkind of dresses, and hotels of a particular star category are a subkind of hotels. That would appear to indicate that features of similarity comparison correspond to properties that serve as modifiers of kinds.

3.2 Internal vs. external modification

The findings above lead us to the distinction between internal and external modification, which was originally suggested by Maienborn (2003) to account for the interpretation of locative event modifiers. In *Anna prepared the chicken in the garden* the modifier locates the whole event of chicken-preparing: it took place in the garden. This is called external modification. In contrast, in *Anna prepared the chicken in the wok* modification is internal expressing a manner of performing the event. The whole chicken-preparing event did not take place in the wok, but instead it was done using a wok. Internal event modifiers specify "some internal aspect of the verb's event argument, whose exact role is left semantically implicit and can only be determined when taking into account conceptual knowledge about the respective event type" (Maienborn and Schäfer 2011, 1411)

In Maienborn, Gese and Stolterfoht (2016) this idea was generalized such that internal event modifiers – including manner adverbs as well as internally used locative and temporal adverbials – pertain to the event kind (thereby extending the account of event kind modifiers in Anderson & Morzycki 2015). In contrast, external event modifiers pertain to the event token. This is captured in

syntax by assuming that internal adverbials adjoin below VP-level whereas external ones adjoin above VP-level. The syntactic localization is supported by positional effects: Internal modifiers have their base position closer to the verb than external ones, see the example in (21b) below.

In Bücking (2012) the external – internal distinction is adapted for nominal locative modifiers, in particular genitive attributes of event nominals, showing that their syntactic position is analogous to the verbal case: internal locatives adjoin closer to the nominal head than external ones, which is mirrored in ordering preferences, see (20) (adapted from Bücking).

- (20) die Zubereitung des Huhns in Pfeffertunke im Garten / ?? im Garten in Pfeffertunke
'preparing the chicken in a pepper marinade in the garden'

Generalizing the ideas in Bücking (2012) and Maienborn, Gese and Stolterfoht (2016) we suggest that the internal-external distinction may apply to nominal as well as verbal modification in general. Internal modifiers are close to the nominal or verbal head adjoining below NP or VP, while external modifiers adjoin above NP or VP, respectively, thereby predicting ordering preferences; see a nominal and a verbal example in (21) (a) and (b).

- (21) a. eine Bluse aus Seide mit Fettflecken / ? eine Bluse mit Fettflecken aus Seide
'a blouse made of silk with grease stains'
b. ... weil Paul das Huhn im Garten im Wok / ?? im Wok im Garten zubereitet hat
'... because Paul cooked the chicken in the garden in a wok'

Furthermore, we follow Carlson (2003) in assuming that the lexical projection of a major phrase contains only type information, whereas in higher functional projections only token information is available. This implies that internal modifiers pertain to types, or kinds, while external modifiers pertain to tokens (as suggested for verbal modifiers by Maienborn, Gese and Stolterfoht 2016). A similar idea can be found in McNally & Boleda (2004) who observe that relational adjectives (in the sense of Bally 1944)¹² always occur closer to the head noun than other intersective adjectives, see (22) (McNally 2017, ex. 8).

- (22) a. a clever legal adviser
b. ?? a legal clever adviser

McNally & Boleda (2004) argue that Bally's relational adjectives denote descriptions of kinds rather than of token entities. Continuing this line of thought, McNally (2017) proposes to distinguish two ways of modification: kind-level modifiers serve to create subkind descriptions while token-level modifiers provide additional descriptions of referents. This idea matches perfectly with the internal-external distinction discussed above.¹³

The internal-external distinction – reflecting kind-level vs. token-level modification – suggests itself as a hypothesis to capture the preliminary findings described in the previous section: Nominal as well as verbal similarity phrases pick up internal modifiers, while external modifiers are dispreferred. Thus, in the nominal case, acceptable antecedents contain modifiers affecting the nominal kind or concept – *bike with gears*, *new iPhone version* – while modifiers expressing a relation between the referent and some other entity – *dress belonging to Aunt Anna*, *laptop offered by Paul* – as well as subjective (judge-

¹²Bally's relational adjectives are semantically similar to nouns conveying the intuition that they are 'covert nouns'; one of his examples is *solaire* as in *chaleur solaire* 'solar heat'.

¹³In (2017) McNally extends her proposal by using distributional representations instead of Carlsonian kind-descriptions (details would go too far here).

dependent) predicates – *beautiful hotel* – are dispreferred. In the verbal case, acceptable antecedents are given by event-internal locatives and genuine manners – *prepare a chicken in the wok / low-fat style* – while event-external locatives – *in the garden* – as well as subjective predicates – *prepare a chicken tastily* – are again dispreferred.

3.3 Experimental study: accessibility of antecedent modifiers by *so*

The hypothesis that internal modifiers of adnominal and adverbial *so* are preferred over external modifiers was tested in an experimental study comparing the acceptability of two-sentence sequences of the form shown in (23) and (24). In the first sentence, the intended property to be picked up was introduced as a nominal or adverbial modifier, and in the second sentence *so* occurred in its anaphoric use. The modifier was added in a *nämlich* 'that is' continuation to make sure that participants end up with the intended interpretation. (Otherwise they might assume that *so* is part of an equative with an elided standard of comparison – *so eine Bluse [wie die von Maria]* 'a blouse like Maria's' – which would make all test items fully acceptable.)

We used only PPs as internal and external modifiers, in order to keep items comparable with regard to syntactic structure and complexity. We predicted that internal modifiers are more accessible by *so* than external modifiers – adnominal as well as adverbial ones.

- (23) a. Maren hat gestern einen Rock aus Wildseide getragen. Angelika hat auch so einen Rock getragen. (... nämlich einen aus Wildseide).
 b. Maren hat gestern einen Rock von ihrer Schwester getragen. Angelika hat auch so einen Rock getragen. (... nämlich einen von ihrer Schwester).
 'Yesterday, Maren wore a skirt made of wild silk / from her sister. Angelika wore a skirt like this, too (... that is, one made of wild silk / from her sister).'
- (24) a. Nora hat ihr Karnevalskostüm mit der Hand genäht. Mandy hat ihr Karnevalskostüm auch so genäht. (... nämlich mit der Hand).
 b. Nora hat ihr Karnevalskostüm in der Schule genäht. Mandy hat ihr Karnevalskostüm auch so genäht. (... nämlich in der Schule).
 'Nora sewed her costume for carnival by hand / in school. Mandy sewed her costume for carnival like this, too (... that is, by hand / in school).'

Materials consisted of 32 experimental sentence pairs (16 with adnominal modifiers, 16 with adverbial modifiers) and 35 filler sentence pairs. Each experimental item was prepared in two versions that differed with respect to the SEMANTIC TYPE of the adverbial in the first sentence (*internal* vs. *external*). The factor SYNTACTIC TYPE (*adnominal* vs. *adverbial*) was realized between items. Two presentation lists were constructed in which the 32 experimental items were randomly mixed with the 35 filler items¹⁴. The two lists were counterbalanced across items and conditions: Each list included only one of the two version of each experimental sentence. 24 participants, all students of Tübingen University, were asked to rate the sentence pairs on a 5 - 1 Likert scale (5 = perfectly acceptable, 1 = completely unacceptable).

The descriptive results are presented in Figure 1. A repeated measures ANOVA revealed a highly significant main effect of SEMANTIC TYPE ($F(1,23) = 27.67, p \leq .0001$; $F(1,15) = 83.88, p \leq .0001$) with higher ratings for the sentence pairs with internal modifiers compared to the external ones. Furthermore, we found a highly significant effect for SYNTACTIC TYPE with higher ratings for nominal items compared to verbal one ($F(1,23) = 26.99, p \leq .0001$; $F(1,15) = 21.24, p \leq .0001$). Since we used

¹⁴ All materials used in the experiment can be found here:

https://osf.io/7p8hm/overview?view_only=0b68a1ddf2ce41ee8b1703594b3dafa2

different lexical materials for nominal and verbal stimuli, this effect cannot be interpreted.¹⁵ Finally, there was no interaction between SYNTACTIC TYPE and SEMANTIC TYPE ($F_1(1,23) = 2.37, p = .14$; $F_2(1,15) = 1.19, p = .29$) suggesting that modifiers in the nominal and the verbal domain behave similar with regard to anaphoric *so*.

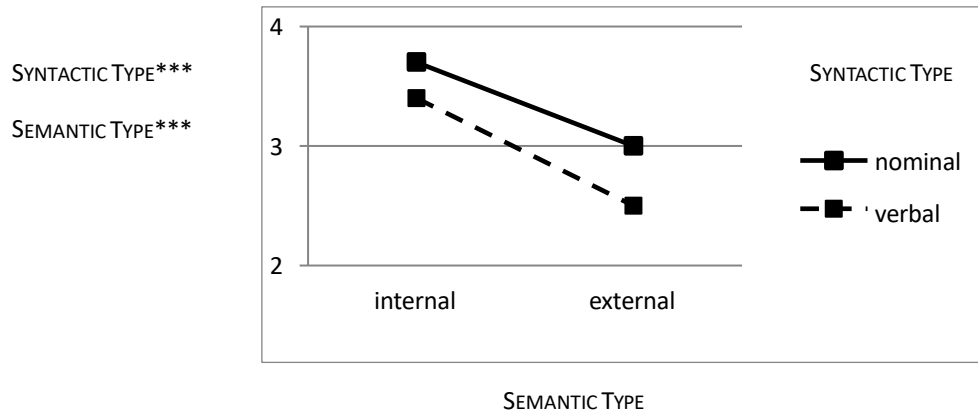


Figure 1. Mean ratings on Likert Scale 5-1

The results support our hypothesis concerning the internal-external distinction, which we assume reflects the difference between kind-level and token-level modification: Internal modifiers as antecedents of adnominal as well as adverbial *so* are preferred in comparison to their external counterparts.

3.4 Intermediate summary

This section started from the observation that there are restrictions on the use of the similarity demonstrative *so* such that some properties are permissible in similarity comparison while others are not, thereby raising the question of what these restrictions look like and, even more importantly, what the reasons for these restrictions are. We considered results of exploratory pretests, all of them pointing to the distinction between internal and external modifiers. While originally applied to locative adverbial modifiers, this distinction can plausibly be adapted to adnominal as well as other types of adverbial modifiers. This led to the hypothesis that adnominal as well as adverbial *so* preferably picks up internal modifiers as compared to external ones, which was confirmed in the experiment.

The internal-external distinction has a syntactic aspect: the modifier is adjoined below or above NP and VP, *rsp.* This accounts for positional effects, among other things. However, there is a semantic distinction underneath which is that between modification of kinds and tokens, *rsp.* The experiment thus showed that kind modifiers are preferred over token modifiers.

This leaves open the question of why certain properties can be understood as kind modifiers and others cannot. Why does *from Nicaragua* have a significantly higher capacity to form a subkind of coffee than *from the canteen*, whereas it has little chance to form a subkind of bikes? This question addresses the nature of the features of comparison from a cognitive point of view, which is the topic of the next section.

¹⁵ According to our native speaker intuition the continuations with *so* modifying the NP sound more common than the VP cases. One reason for this difference could be the complexity of the referent (entity vs. event).

4 Similarity and kind formation

This section starts with the question of what enables certain properties to act as internal modifiers. Prasada and colleagues provide experimental evidence that there are properties principally connected to kinds – we will refer to them as *in-virtue-of* properties. There is reason to assume that these properties correspond to internal modifiers because they both pertain to kinds, which will turn out to be true even if not directly. Since internal modifiers have been shown to correspond to features permissible in similarity comparison, there is, moreover, reason to assume that the latter also pertain to kinds, and thus to hypothesize that similarity phrases establish ad-hoc subkinds.

4.1 Principled vs. statistical connections between properties and kinds

In a number of papers, Prasada and colleagues report on experiments investigating generic knowledge about kinds (see Prasada & Dillingham 2006, Prasada et al. 2012, Prasada et al. 2013, Haward et al. 2021). Evidence is provided that kinds may be connected to properties in (at least) two different ways. There may be a *principled connection* or a *statistical connection* between a kind and a property. Principally connected properties – *k-properties* – apply to objects of a kind in virtue of being that kind: A dog has four legs in virtue of being a dog. This has a normative flavor suggesting that a dog should have four legs. If it has three legs, there is something wrong, maybe it had an accident. In contrast, statistically connected properties – *t-properties* – are just prevalent: Most dogs wear collars, but this cannot be understood as applying to dogs by virtue of them being dogs, and there is nothing wrong with a dog that does not wear a collar.¹⁶

The distinction between k- and t-properties was confirmed in a number of experiments testing paraphrases (*by virtue of* vs. *in general*), formal explanations (*why does that (pointing to a dog) have four legs? Because it is a dog.*) and normativity (*dogs should have four legs*). Finally, the distinction between k- and t-properties is reflected in generic sentences: Bare plurals are fine with both k- and t-properties, while indefinite singulars are acceptable with k-properties only, see (25c-c").¹⁷

- (25)
- | | |
|-----|---|
| a. | Dogs are four-legged by virtue of being the kinds of things they are. |
| a' | # Dogs wear collars by virtue of being the kinds of things they are. |
| b. | Dogs in general are four-legged. |
| b' | Dogs in general wear collars. |
| c. | A dog / dogs have four legs. |
| c'. | # A dog wears a collar. |
| c". | Dogs wear collars. |

Prasada and Dillingham (2006) discuss a general structural constraint on k-properties in kind taxonomies: "If a property p is represented as a k-property of a given kind K, then there can be no superordinate kind K' of K, for which p is also a k-property." (p. 105). This constraint is motivated by the observation that (26a) is acceptable while (26b) is not, thereby suggesting that being extended in three

¹⁶ An explanation in terms of the kind something belongs to is called the *formal mode* of explanation and is one of the four modes of explanation described by Aristotle (see Prasada 2017). From the point of view of a teleological mode, instead of the formal mode, one might explain the fact that dogs are four-legged by their need to be fast runners.

¹⁷ This observation is in accordance with the definitional nature of indefinite singular generics suggested by Krifka (2012).

dimensions is a k-property of the kind MATERIAL BEING, not of the kind DOG. Likewise, being four-legged is a k-property of DOG, not of DACHSHUND, see (27a, b).

- (26) a. Dogs, by virtue of being material beings, are extended in three dimensions.
b. #Dogs, by virtue of being dogs, are extended in three dimensions.
- (27) a. Dachshunds, by virtue of being dogs, have four legs.
b. #Dachshunds, by virtue of being dachshunds, have four legs.

This raises the question, however, which superordinate kinds are under discussion. There are a number of kind terms qualifying for being principally connected to being four-legged (*Canis lupus*, quadruped, four-legged animal). Moreover, the lack of acceptability of (27b) seems to suggest that being four-legged is a t-property of dachshunds which is clearly wrong – they don't have four legs by mere accident. Still, they don't have four legs by being dachshunds either, but inherit this property from a superordinate kind (DOG, CANIS LUPUS, ..). Thus, the question of whether a certain property is a k-property of a kind K depends on the way K is conceived in a given taxonomy of kinds.

4.2 Do internal modifiers express principled connections?

In *Anna has an electric car* the modifier *electric* is clearly internal (and can be picked up by *such* – *Berta has such a car, too.*) Since it modifies a kind, there is reason to assume that *electric* is a k-property – but of which kind? If it were the kind ELECTRIC CAR, we would end up with the by-virtue-of statement in (28a) which sounds awfully redundant. But if it were the kind CAR, the by-virtue-of statement would be plainly false, (28b).

- (28) a. ?? Electric cars are electric by virtue of being electric cars.
b. # Electric cars are electric by virtue of being cars.

At the same time, there is a strong intuition that *electric* is connected to the kind CAR in a non-accidental way. This is apparent when comparing *electric car* to *registered car*. Being registered is clearly connected to the kind of cars but not in a principled way – a car need not be registered (apart from the point of view of the tax administration). A car does not have to be electric either, but it has to have some type of a drive in order to count as a car. Thus, there is a principled connection between the kind CAR and the feature *have an x-type of drive*, where x-type is one of *electric* or *gasoline* or *diesel* etc. It is important to note that this is a feature in the strict sense rather than a property, and has to be combined with a value to yield a property (see section 2.3). The by-virtue-of statement thus has to look like a disjunction or a modal statement, see (29) (a), (b).

- (29) a. Cars have a drive by virtue of being cars. (where a drive is electric or gasoline etc.)
b. Cars may have an electric drive.

So there is a principled – by-virtue-of – connection, but it is not between a kind and a property but instead between a kind and a feature (or dimension, if you prefer). We will call it a *k-feature*. Coming back to the question in the title of this section, the answer is 'No, but yes'. Internal modifiers do not correlate to k-properties but provide values (e.g., *electric* or *gasoline* or *diesel*) for k-features (e.g., *have some type of a drive*), which are principally connected to a kind and thereby have to be in-built in the representation of a kind.

4.3 Do similarity phrases correspond to kinds?

From the point of view of the similarity analysis, similarity demonstratives denote properties: be similar to the item pointed at. Being modifiers, we would first of all expect them to combine via intersective modification with the denotation of their head noun – when pointing to three-legged tables, *solche*

Tische 'such tables' would be a subset of tables that are three-legged. We saw, however, that the features permissible in similarity comparison are confined to those that are principally connected to the kind denoted by the head noun. This is reason to hypothesize that similarity phrases establish ad-hoc subkinds instead of mere subsets.

We explored this hypothesis with the help of kind-level predicates. This type of predicates can be applied to kinds, but not to individual items. Thus compatibility with a noun is clear evidence for a kind interpretation. Standard examples are *ausgestorben* / *extinct* and *weitverbreitet* / *widespread* (see Krifka et al. 1995, section 1.3). Additional examples would be *erfunden worden sein* / *be invented, in Mode sein* / *be stylish* and *sich zum Verkaufsschlager entwickeln* / *become a bestseller*.

Since indefinite singulars are not suited to be tested with kind-level predicates,¹⁸ we use plural similarity phrases like *solche Mäuse* 'such mice'. In (30) – (32) examples from the experiment and from pretests are modified using plural similarity expressions plus kind-level predicates. For example, (30a) is a modified version of (11a) in which *Solche Räder haben sich zum Verkaufsschlager entwickelt* 'such bikes became a big seller' has been substituted for *Berta hat auch so ein Rad* 'Berta has a bike like that, too.'. The original versions of (30)-(32) (a) are acceptable and thus, according to our hypothesis, the modified versions should be acceptable. The original versions of the (b) cases are unacceptable and the modified ones should also be unacceptable. This is, in fact, what we find.¹⁹

¹⁸ Indefinite singulars are never compatible with kind-level predicates regardless of what they denote, which is considered as evidence that they never refer to kinds: # *A striped dormouse is extinct* (see Krifka et al. 1995)

¹⁹ We also explored the hypothesis that similarity phrases establish ad-hoc subkinds by considering paraphrases based on the German term *Art* ('kind/species'). For example, in the sequence in (31a) *solche Röcke* denotes the kind 'skirts made of wild silk'. This is demonstrated by successful combination with the kind-level predicate *in Mode sein* 'be in fashion'. This would suggest that similarity phrases picking up permissible features of comparison allow for an *Art* paraphrase of the ad-hoc subkind, as in (a), while with non-permissible features the *Art* paraphrase fails.

- (33) a. Röcke aus Wildseide sind eine Art von Röcken.
 b. # Röcke von Marens Schwester sind eine Art von Röcken.
 'Skirts made of wild silk / from Marens sister are a kind of skirts.'

However, this line of reasoning overlooks the fact that the German term *Art* need not coincide —and in fact does not coincide —with the technical term 'kind', as used in semantics. We found that if an *Art* paraphrase is acceptable the corresponding similarity phrase denotes a kind. But we also found cases in which an *Art* paraphrase is not acceptable even though the similarity phrase does denote a kind. In the sequence in (32a) and (32b), electric toothbrushes are shown to be a kind whereas disgusting toothbrushes are not. This is reflected by the (un)acceptability of the *Art* paraphrase in (34a) and (b). However, in combination with *zerkaut* 'chewed-up' the similarity phrase denotes a kind, as evidenced by combination with the kind-level predicate *weitverbreitet* 'widespread' in (32c), but the *Art* paraphrase in (34c) clearly fails.

- (32) c. Auf der Ablage über dem Waschbecken liegt eine zerkaute Zahnbürste. Solche Zahnbürsten sind leider weitverbreitet (nämlich zerkaut).
 'There is a chewed-up toothbrush on the shelf above the sink. Unfortunately, toothbrushes like this are widespread (that is, chewed-up ones).'
- (34) a. Eine elektrische Zahnbürste ist eine Art von Zahnbürste.
 b. Eine eklige Zahnbürste ist eine Art von Zahnbürste.
 c. # Eine zerkaute Zahnbürste ist eine Art von Zahnbürste.
 'An electric / disgusting / chewed-up toothbrush is a kind of toothbrush.'

Prasada et al. (2012) consider a related idea. They hypothesize that – in English – simple nouns as well as compounds but not nouns combined with adjectives denote kinds. The hypothesis was successfully tested by paraphrases with English *kind* – *sunflowers are a kind of flowers*. The failure of the *Art*-paraphrase in testing for kindhood shows, first, that the meaning of the German term *Art* is not equivalent to the English term *kind*, which is no surprise given that *type nouns*, as they are called by Brems et al. (2022), vary considerably in their coverage within and across languages, see Umbach (2022). Secondly, "trying to identify lexical items as theoretical terms typically leads to dead ends" as one reviewer pointed out – we agree, thank you.

- (30) a. Anna hat ein Rad mit Gangschaltung. Solche Räder haben sich zum Verkaufsschlager entwickelt (nämlich Räder mit Gangschaltung).
 b. Anna hat ein neues Rad. # Solche Räder haben sich zum Verkaufsschlager entwickelt. (nämlich neue).
 'Anna has a bike with gears / a new bike. Such bikes became a big seller (that is, bikes with gears / new bikes).'
- (31) a. Maren hat gestern einen Rock aus Wildseide getragen. Solche Röcke sind jetzt in Mode (... Röcke aus Wildseide).
 b. Maren hat gestern einen Rock von ihrer Schwester getragen. # Solche Röcke sind jetzt in Mode (... nämlich Röcke von Maren's Schwester).
 'Yesterday, Maren wore a skirt made of wild silk / from her sister. Such skirts are in fashion now (...that is, skirts made of wild silk / from her sister).'
- (32) a. Auf der Ablage über dem Waschbecken liegt eine elektrische Zahnbürste. Solche Zahnbürsten sind weitverbreitet (... nämlich elektrische).
 b. Auf der Ablage über dem Waschbecken liegt eine eklige Zahnbürste. # Solche Zahnbürsten sind weitverbreitet (... nämlich eklige).
 'There is an electric / disgusting toothbrush on the shelf above the sink. Toothbrushes like this are widespread (... that is, electric ones / disgusting ones).'

In the example in (35) *so* is used adverbially. The original version is in (24). Testing event types for kindhood with the help of kind-level predicates has not been sufficiently researched. Still, the example in (35) suggests that verbal similarity phrases constitute kinds – event kinds – as do their nominal counterparts.

- (35) a. Nora hat ihr Karnevalskostüm mit der Hand genäht. Ein Karnevalskostüm *so* zu nähen ist weitverbreitet (... nämlich mit der Hand).
 b. Nora hat ihr Karnevalskostüm in der Schule genäht. # Ein Karnevalskostüm *so* zu nähen ist weitverbreitet (... nämlich in der Schule).
 'Nora sewed her costume for carnival by hand / in school. Sew a costume for carnival like this is widespread (... that is, by hand / in school).'

Summing up, similarity phrases do in fact correspond to kinds. They denote subkinds of the kind denoted by the head noun or verb, and they are ad-hoc because the features picked up in the similarity phrase are not given a priori. On the other hand, ad-hoc kinds cannot be formed at will. Ad-hoc kind formation is constrained by the range of features permissible in similarity comparison, which have to be k-features, i.e. be principally connected to the kind denoted by the noun or verb.

In the end, this result confirms the idea of Carlson (1977) and subsequently Anderson and Morzcki (2015) that phrases headed by English *such* refer to kinds. But now we know how this connection comes about: Kind reference of these phrases is due, first, to the similarity expressed by *such* and, secondly, to the restriction of the features of similarity comparison to principally connected ones.

5 Ad-hoc kind formation beyond similarity phrases

Similarity phrases correspond to kinds because features of similarity comparison must be principally connected to the kind denoted by the head noun (or verb). Since these kinds do not correspond to

lexical expressions we call them *ad-hoc kinds*. An obvious question then is whether there are other variants of ad-hoc kinds and what strategies they are based on: Is ad-hoc formation of kinds necessarily based on similarity, or are there other conceivable strategies?

The notion of ad-hoc kinds is, first of all, reminiscent of Barsalou's notion of *ad-hoc categories* (Barsalou 1997, 2010). In a series of experiments, Barsalou investigated the structure of what he termed *ad-hoc categories*. They are created by goal-oriented nominal descriptions, for example '*ways to make friends*' and '*things to take from one's home during a fire*'. In contrast to categories denoted by lexicalized expressions (*dogs, chairs*, etc.), ad-hoc categories are not well-established in memory and may violate the correlational structure of the environment (see Rosch 1978). For example, instances of the category '*things to take from one's home during a fire*' may be as diverse as a box of letters, a computer, a guinea pig, and one's favorite baseball cap. Experimental evidence is provided showing that ad-hoc categories are gradient, as are regular categories – some items are ranked to be better exemplars than others. They differ, however, from regular categories in showing only weak associations between instances and concepts, no more than random nominal descriptions.

Barsalou's notion of categories is not the same as that of kinds, as we use it in this paper. Still, well-established categories are associated with familiar words, and categories, well-established or ad-hoc, are said to index concepts. This is close to the Carlsonian view of kinds as corresponding to labeled concepts (which we adopt here). Barsalou's ad-hoc categories are not collected by similarity but rather by fitness to serve a specific purpose. They do not constitute a subkind of a given kind but may range across different kinds. Thus, Barsalou's ad-hoc categories resemble the ad-hoc kinds examined in this paper only insofar as they are constructed ad-hoc instead of being lexically given.

Mendia (2020) discusses ad-hoc kinds instead of categories, though from a perspective close to Barsalou's assuming that *ad-hoc* kinds may disregard natural regularities. He focuses on modification by relative clauses and argues that modified plural definite descriptions are genuine kind-referring DPs compatible with kind-level predicates: *The lions that eat people are widespread*, (p.594). Restrictions on similarity demonstratives are not recognized as restrictions on subkind formation because the in-virtue-of test is said to fail, without, however, considering a modal version (see our notion of k-features in section 4.2). Mendia concludes that there are "two closely related but ultimately different ways of referencing ad-hoc kinds" (p. 607).

Barsalou's idea of ad-hoc categories is adopted in Mauri (2017) and subsequent works. Mauri takes a linguistic perspective focusing on strategies of encoding, i.e. overtly signaling, the process of ad-hoc category building across languages. Linguistic constructions examined in her work include special plurals and other derivational forms, as well as open-ended lists and general extenders.

Mauri adopts the research on plurals in Moravcik (2003)/(2020), who distinguishes between associative and similitive plurals (in addition to regular ones), see (36) and (37). While associative plurals turn a single individual into a group including the individual plus associated entities, similitive plurals include entities similar to those denoted by the basic expression, see Krifka and Modarresi (2023). Mauri, moreover, points out that in Italian there is a derivational strategy close in meaning to associative plurals where the affix *-ame* is attached to animate as well as inanimate roots deriving collective nouns, see (38). Finally, Moravcik in her concise paper in (2020) on "The place of ad-hoc categories within the typology of plural expressions" demonstrates the difference between similitive and associative plurals with the help of the example in (39).

- | | | | |
|------|---------------------|---|--------------------------------------|
| (36) | <i>Peter-ek</i> | 'Peter and his family or friends or associates' | Hungarian, Moravcik (2003) |
| (37) | <i>ketāb 'book'</i> | <i>ketāb metāb</i> | 'book(s) and the like' |
| | | | Persian
Krifka & Modarresi (2023) |

- (38) *scatolame* 'set of boxes of various types' Mauri (2017), 311
berlusconame 'people having to do with Berlusconi' or 'Berlusconi and everything that goes with it' (including parties, luxury dinners, corruption, etc.)
- (39) a. Plurality based on similarity: *bread, cake, and muffins* Moravcik (2020), ex. (7)
 b. Plurality based on grouphood: *bread, butter, and cheese*

The other strategy of forming ad-hoc categories Mauri focuses on is that of open-ended lists supplemented by general extenders. Open-ended lists are coordinations with an open ended interpretation usually indicated by a general extender expression like *etc.*, *and the rest*, or *and such*, as in (40), see Mauri (2017). The idea that open-ended lists form ad-hoc categories is reminiscent of Lang's (1984) coordination conditions. Lang showed that, from the perspective of semantics, coordinated elements have to be semantically independent – neither of them subsuming the other – and secondly, there has to be a "common integrator", i.e. a category or concept subsuming the conjuncts, the nature of which has been under debate. In Umbach (2004) it is argued that the "common integrator" is formed by similarity of the conjuncts. From this point of view, ad-hoc categories formed by open-ended lists make use of similarity, which is which is often made explicit by a supplement expressing similarity (*and the like*).

- (40) *water, herbal teas, smoothies, and the like*

Coming back to the question whether strategies of generating ad-hoc categories are generally based on similarity: open-ended lists as well as similative plurals clearly are whereas associate plurals are not even though grouphood requires some homogeneity – the group in (36) may consist of Peter and his friends, but not of Peter and his dogs (see Moravcik 2003). The next question would be whether similative plurals and open-ended lists denote kinds or just plural individuals, and are subject to the constraints discussed in this paper. This question has to be left for future work.

6 Conclusion

This paper starts from the analysis of similarity demonstratives in Umbach and Gust (2014) – German *so/solch*, English *such*, Polish *tak*, etc. – arguing that they express similarity (instead of identity) to the target of the demonstration. Similarity is spelt out as indistinguishability with respect to a given set of features of comparison. The focus of the paper is on restrictions on features permissible in similarity comparison. Experimental evidence was presented showing that these restrictions correspond to the difference between internal and external modification: Internal – kind related – modifiers are permissible in similarity comparison, while external – referent related – modifiers are dispreferred.

This result led to the question what enables certain properties but not others to act as kind modifiers. We considered the work by Prasada and colleagues, who argue that some properties are principally connected to the kind they are predicated of, while others are not even if they are prevalent in objects of this kind. It turned out that permissible features of similarity comparison correspond to principally connected properties (though in a modalized version). Building on this, evidence was provided that similarity phrases establish ad-hoc subkinds of the kind denoted by the head noun or verb.

It has been argued before by Carlson (1977) and subsequently by Anderson and Morzcki (2015) that phrases headed by English *such* and Polish *tak*, *rsp.* refer to kinds, without, however, asking the question how these kinds come about, and without trying to explain the obvious restrictions. The similarity account provides a transparent analysis explaining why similarity phrases refer to kinds and, moreover, offers a way to explain the observed restrictions.

The question underlying the entire paper is: what is the nature of the features permissible in similarity comparison? Maybe they reflect a more general characteristic found in all phenomena we have examined: internal modification as described in Maienborn (2003) and Bücking (2012), relational adjectives considered in McNally & Boleda (2004), and finally the principally connected properties studied by Prasada and colleagues (e.g. 2006, 2013).

One way to approach this characteristic might be through the notion of *intrinsic property* as discussed in philosophy: "A property is intrinsic iff individuals have it in virtue of how they themselves are, not in virtue of the relations in which they stand to other individuals." (Hoffmann-Kolss 2024). A famous test illuminating the basic intuition about the intrinsic/extrinsic distinction is the so-called loneliness test, according to which an intrinsic property can be held by individuals independently of whether they are "lonely" or accompanied by other (distinct) individuals (see Langton and Lewis 1998 and the discussion in, e.g., Hoffmann-Kolss). This is, of course, a metaphysical perspective, whereas this paper takes a semantic approach, with a detour into cognitive science. Nevertheless, it appears tempting to think about intrinsicity, since external modifiers (that are not permissible in similarity comparison) seem to systematically violate the loneliness test: A "lonely" table may still have three legs but not be owned by Aunt Elsi because there is no Aunt Els), and it may not be nice because subjective properties require judges thereby failing loneliness. To what extent these considerations are viable is a question for future work.

References

- Anderson, C., and M. Morzycki. 2015. Degrees as kinds. *Natural Language and Linguistic Theory*. 33–3, 791–821.
- Bally, Charles. 1944. *Linguistique générale et linguistique française*. A. Francke, Berne.
- Barsalou, L. W. 1983. Ad-hoc categories. *Memory & Cognition* 11 : 211– 27.
- Brems, L., Davidse, K., Hennecke, I., Kolyaseva, A., Kisiel, A. & Mihatsch, W. 2022. (Eds.) *Type Noun Constructions in Slavic, Germanic and Romance Languages: Semantics and Pragmatics on the Move*. De Gruyter Mouton..
- Bücking, S. 2012. Kompositional flexibel. Partizipanten und Modifikatoren in der Nominaldomäne. Tübingen: Stauffenburg.
- Carlson, G. 1977. *Reference to kinds in english*. Ph.D. dissertation, University of Massachusetts, Amherst. Published 1980 by Garland Press, New York.
- Carlson, G. 2003. Interpretive Asymmetries in Major Phrases. In A-M. DiSciullo (ed.) *Asymmetry in Grammar*. Benjamins. 299-311.
- Carlson, G. 2010. Generics and concepts. In F. J. Pelletier (ed.) *Kinds, Things and Stuff*. Oxford, Oxford University Press, 16-36.
- Fishman, Alon. 2020. English similarity predicates construe particular dimensions of similarity. *Cognitive Linguistics* 31(3), 453-484.
- Gärdenfors, P. 2000. *Conceptual Spaces*. MIT Press.
- Goodman, N. 1972. Seven strictures on similarity. In Goodman, N. (ed.) *Problems and Projects*. The Bobbs Merrill Company, Indianapolis and New York, 437–447.
- Greenberg, Y. 2007. Exceptions to Generics : Where Vagueness, Context Dependence and Modality interact. *Journal of Semantic* 24.
- Haward, P., S. Carey, & S. Prasada. 2021. The Formal Structure of Kind Representations. *Cognitive Science* 45 (10).
- Hoffmann-Kolss, V. 2024. Intrinsic/Extrinsic. In Fisher, A.J.R., Maurin, A.-S. (eds.), *The Routledge Handbook of Properties*, Routledge: 92-102.
- Kaplan, D. 1989. *Demonstratives*. In J. Almog, J. Perry & H. Wittstein (eds.), *Themes from Kaplan*, Oxford University Press, 481–563.

- Klein, E. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4(1), 1–45.
- König, Ekkehard & Carla Umbach (2018) Demonstratives of Manner, of Quality and of Degree: A Neglected Subclass. In M. Coniglio, A. Murphy, E. Schlachter & T. Veenstra (eds.), *Atypical demonstratives: syntax, semantics and pragmatics*. Berlin, de Gruyter Mouton.
- Krifka, M., Pelletier, F. J., Carlson, G., ter Meulen, A., Link, G. & Chierchia, G. 1995. Genericity: An introduction. In Carlson, G. & Pelletier, F. J. (eds.) *The Generic Book*. Chicago: University of Chicago. 1-124.
- Krifka, M. 2012. Definitional generics. In A. Mari, C. Beyssade & F. Del Prete (eds.), *Genericity*. Oxford University Press. Oxford. 372–89.
- Krifka, M. & F. Modarresi. 2023. Similitive Plurals in Persian: Alternatives, Homogeneity, or Strongest Meanings? *Proceedings of the 41st West Coast Conference on Formal Linguistics*, University of California.
- Lang, E. 1984. *The semantics of coordination*. SLCS 9, Amsterdam, Benjamins.
- Langton, R. & Lewis, D. 1998. Defining "Intrinsic". *Philosophy and Phenomenological Research* 58(2): 333-345.
- Maienborn, C. 2003. Event-internal adverbials: Semantic underspecification and conceptual interpretation. In: E. Lang, C. Maienborn & C. Fabricius-Hansen (eds.) *Modifying Adjuncts*. Berlin: de Gruyter, 475–509.
- Maienborn, C. & M. Schäfer. 2011. Adverbs and adverbials. In K. von Stechow, C. Maienborn & P. Portner (eds.) *Semantics* (HSK 33.2), de Gruyter, 1390–1420
- Maienborn, C., H. Gese & B. Stollerfoht. 2016. Adverbial Modifiers in Adjectival Passives. *Journal of Semantics*, 33, 299–358.
- Mauri, C. 2017. Building and Interpreting Ad-Hoc Categories: A Linguistic Analysis. In J. Blochowiak, C. Grisot, St. Durrleman, Ch. Laenzlinger (eds) *Formal Models in the Study of Language*. Springer, 297-326.
- McNally, L. 2017. Kinds, descriptions of kinds, concepts, and distributions. In K. Balogh & W. Petersen (eds.) *Bridging Formal and Conceptual Semantics*. Selected Papers of BRIDGE-14. Düsseldorf, 39-61.
- McNally, L. & G. Boleda. 2004. Relational adjectives as properties of kinds. In O. Bonami & P. Cabredo Hofherr (eds.) *Empirical Issues in Formal Syntax and Semantics 5*, 179–196.
- Mendia, J. A. 2020. Reference to *ad-hoc* kinds. *Linguistics and Philosophy* 43:589–631
- Moravcsik, E. 2003. A semantic analysis of associative plurals. *Studies in Language* 27, 3, 469 – 503.
- Moravcsik, E. 2020. The place of ad-hoc categories within the typology of plural expressions. *Language Sciences*. 81.
- Nunberg, G. 1993. Indexicality and Deixis. *Linguistics and Philosophy* 16: 1-43.
- Prasada, S., & E. M. Dillingham. 2006. "Principled and Statistical Connections in Common Sense Conception." *Cognition* 99 (1): 73–112.
- Prasada, S., L. Hennefield, & D. Otap. 2012. Conceptual and Linguistic Representations of Kinds and Classes. *Cognitive Science* 36 (7): 1224–1250.
- Prasada, S., S. Khemlani, S.-J. Leslie, and S. Glucksberg. 2013. "Conceptual Distinctions Amongst Generics." *Cognition* 126 (3): 405–422. <https://doi.org/10.1016/j.cognition.2012.11.010>.
- Prasada, S. 2017. The scope of formal explanation. *Psychonomic Bulletin & Review*.
- Prasada, S., and D. G. Hall. 2019. Instance-of-object-kind Representations. *Cognition* 189 (C): 209–220.
- Rosch, E. 1978. Principles of Categorization. In E. Rosch and B. Lloyd (eds.) *Cognition and Categorization*. Laurence Earlbaum Associates, Hillsdale, 27-48.
- Sassoon, G. 2013. A typology of multidimensional adjectives. *Journal of Semantics*, 30, 335–380.

- Specht, L. & Stolterfoht, B. 2022. Incrementality in the processing of adverbial order variations in German. In S. Featherston, R. Hörnig, A. Konietzko & S. von Wietersheim (Eds.), *Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data* (pp. 135-151). Tübingen: University of Tübingen. <http://dx.doi.org/10.15496/publikation-75911>
- Tversky, A. 1977. Features of similarity. *Psychological Review* 84: 327-352.
- Umbach, C. 2004. On the Notion of Contrast in Information Structure and Discourse Structure. *Journal of Semantics*, Volume 21, Issue 2, 1 - 21.
- Umbach, C. (2022) Ways of classification: German *Art* and *Typ* . In Brems et al.
- Umbach, C. & H. Gust (2014) Similarity Demonstratives. *Lingua* 149, 74-93.
- Umbach, C. & H. Gust (2021) Grading similarity. In Löbner, S., Gamerschlag, T., Kalenscher, T., Schrenk, M., Zeevat, H. (Eds.) *Concepts, Frames and Cascades in Semantics, Cognition and Ontology*, Springer, 365-388.