Umbach, Carla & Helmar Gust (2014) Similarity Demonstratives. Lingua 149, 74-93

Similarity Demonstratives[1]

Carla Umbach
(Zentrum für Allgemeine Sprachwissenschaft Berlin)

Helmar Gust
(Institut für Kognitionswissenschaft, Universität Osnabrück)

Final draft, April 2014

Similarity demonstratives are a particular class of demonstratives found across languages, e.g., German *so*, English *such*, Polish *tak* and Turkish *böyle*. These demonstratives occur as modifiers of nominal and/or verbal and adjectival phrases posing the problem of how to reconcile their demonstrative characteristics with their modifying capacity. It is hypothesized in this paper that they express similarity to – instead of identity with – the target of the demonstration gesture thereby generating ad-hoc kinds. The paper focuses on the interpretation of ad-nominal and ad-adjectival occurrences of German *so*. An analysis is presented combining truth-conditional semantics with multi-dimensional attribute spaces with the help of a generalization of adjectival measure functions. Similarity is defined in such spaces as indistinguishability with respect to a given set of relevant dimensions.

Similarity; demonstratives; ad-hoc kinds; multi-dimensional attribute spaces; generalized measure functions; indistinguishability.

1 Introduction

In Cognitive Science, similarity is a basic concept of human cognition explaining cognitive skills like perception, classification and learning. Following Tversky, "similarity […] serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations." (Tversky 1977, p. 327). Quine likewise argues that "… there is nothing more basic to thought and language than our sense of similarity; our sorting of things into kinds." (Quine 1969, p. 116). Considering its role in human cognition it is no surprise that there are multiple ways to express similarity in natural languages, for example by lexical items such as German *ähnlich/ gleich/ dasselbe* and English *similar/ like/ the same* and by comparison constructions like equatives and similes. This paper focuses on German *so* ('such'/'like this'). Similar to Dutch *zo*, Polish *tak* and Turkish *böyle* it is a demonstrative occurring as a

[1]This article benefitted substantially from three extremely constructive and sophisticated reviews. We would like to express our gratitude to the anonymous reviewers and to the editors of the special issue, Berit Gehrke and Elena Castroviejo. Valuable comments were provided by Louise McNally, Klaus Robering, Stephanie Solt, Klaus von Heusinger and the audiences of workshops/colloquia in Osnabrück, Madrid, Bielefeld, Tübingen and Sklarska Poreba. Finally, this project has substantially benefitted from the interest and support of Ewald Lang, and from many lively discussions with him on the topic of this paper. We regret that he did not have the opportunity to see the final version. The first author acknowledges financial support by the Deutsche Forschungsgemeinschaft DFG (UM 100/1-1).

modifier in various syntactic phrases thereby raising the question of how to reconcile their demonstrative characteristics with their modifying capacity. We will call this class of demonstratives *similarity demonstratives*.

The German demonstrative *so* modifies adjectival, nominal and verbal phrases and is at the same time a genuine demonstrative expression which can be used deictically and anaphorically. The example in (1) exhibits the deictic use. In (1a) Anna's height is characterized as being similar to the height of the person the speaker points at, in (1b) Anna's car is characterized as being similar in certain respects to the car the speaker points at and finally, in (1c) Anna's way of cutting fish is characterized as being similar to the fish-cutting event the speaker points at.

(1) a. (speaker pointing to a person): So groß ist Anna (auch).
'Anna is this tall, (too).'

b. (speaker pointing to a car in the street): So ein Auto hat Anna (auch).
'Anna has such a car / a car like this, (too).'

c. (speaker pointing to someone preparing a fish): So hat Anna den Fisch (auch) zerlegt.
'Anna cut the fish like this, (too).'

Comparing the demonstrative *so* to standard demonstratives like *dieser*/*this*, the question arises of what the speaker points at when using *so* deictically. Consider (1b): Does he point to a property or kind of cars, or to the actual car in front of him? In standard analyses like Carlson (1980) (for English *such*) and Anderson & Morzycki (2013) (for German *so* and Polish *tak*) it is assumed in cases like (1b) that the speaker points at kinds. It may, however, be doubted whether one can point at kinds and, more importantly, there is empirical evidence against the idea of pointing at kinds when using similarity demonstratives. In this paper, an analysis of similarity demonstratives is proposed such that, first, the speaker points at individuals in (1a/b) and at events in (1c), and secondly, similarity demonstratives differ from standard demonstratives like *dieser*/*this* in expressing similarity between the referent of the demonstrative phrase and the target of the pointing gesture whereas standard demonstratives express identity.

The paper starts from Nunberg's (1993) adaptation of the Kaplanian analysis, interpreting demonstratives as directly referential expressions, but at the same time dismissing the idea that the target of the pointing gesture is necessarily identical to the referent of the demonstrative. This allows for a straightforward interpretation of similarity demonstrative such that the target of the demonstration is the individual, or event, the speaker points at, and the referent of the demonstrative phrase is related to the target by similarity instead of identity. In (1b), for example, it will be assumed that the target of the demonstration is the car the speaker points at and the referent of the NP *so ein Auto* is to be similar, but not identical, to the car pointed at. From this point of view, using similarity demonstratives generates similarity classes – e.g. the class of individuals which are cars and similar to the car pointed at – which can be considered as *ad-hoc* kinds. Thus in contrast to standard analyses like Carlson (1980) and Anderson & Morzycki (2013) the kinds denoted by phrases like *so ein Auto* / *such a car* need not be given in advance and are instead ad-hoc generated by the use of the similarity demonstrative.

For the idea of similarity demonstratives to gain substance, a notion of similarity is required which is not a semantic primitive. It will be proposed to take advantage of the findings on similarity in Artificial Intelligence implementing similarity with the help of multi-dimensional attribute spaces. These spaces are close to Gärdenfors' (2000) conceptual spaces but they provide a qualitative similarity measure instead of a geometrical one. Multi-dimensional attribute spaces will be combined with standard truth-

conditional semantics by means of "generalized measure functions" which generalize the idea of adjectival measure functions (cf. Kennedy 1999) in the following way: While adjectives are one-dimensional associated with a metrical scale, nouns and verbs are multi-dimensional associated with scales of various types – metrical, ordinal, or nominal. So while adjectival measure functions map individuals to degrees on a single dimension, generalized measure functions map individuals to points or regions in multi-dimensional spaces. Combining standard truth-conditional semantics with multi-dimensional spaces will allow for a non-primitive notion of similarity defined as indistinguishability with respect to a given set of relevant dimensions.

In this paper, the idea of similarity demonstratives will be fleshed out for German *so* focusing on the ad-adjectival and ad-nominal occurrences in their deictic use, as in (1a,b). Ad-verbal occurrences, anaphoric uses and expressions of similarity other than *so* will be set aside. In section 2 the deictic use of the demonstrative *so* will be discussed. In section 3, multi-dimensionality of nouns and the idea of generalized measure functions will be presented. In section 4, multi-dimensional attribute spaces will be introduced and a notion of similarity adequate for the interpretation of the demonstrative *so* will be defined.

2 The meaning of the German demonstrative *so*

2.1 Deictic and anaphoric uses of *so*

Although there is a broad range of uses,[2] it is widely agreed that German *so* is in the first place a demonstrative expression used deictically as well as anaphorically. The deictic use has been shown in (1), the demonstrative combining with an adjective in (1a), a noun in (1b) and a verb in (1c). Note that the adjective must be gradable, while the noun and the verb can be either gradable or non-gradable; in combination with gradable expressions, *so* is called *scalar*.[3] The anaphoric use in (2) is parallel to the deictic use: In (2a) Anna's height is characterized as being similar to the height of a previously mentioned person, in (2b) Anna's car is characterized as being similar in certain respects to a previously mentioned car, and in (2c) Anna's way of cutting fish is characterized as being similar to a previously described fish-cutting event.

(2) anaphoric

a. Berta ist 180 groß. Anna ist auch so groß.
   'Berta is 1,80. Anne is that tall, too.'

b. Bertas Auto hat eine Ladeklappe. Anna hat auch so ein Auto.
   'Berta's car has a hatch. Anna has a car like that, too.'

c. Berta zerlegte den Fisch in fünf Teile. Anna hat das auch so gemacht.
   'Berta cut the fish in five parts. Anna did it like that, too.'

The function of the demonstrative in all of the examples in (1) and (2) is that of a modifier of the expression it combines with. In (1a)/(2a) it is a degree modifier, like a measure phrase. In (1c)/(2c) it is a manner modifier, like an adverbial phrase. In (1b)/(2b) the situation is slightly more complicated due to the pre-determiner position of *so* suggesting that it modifies the determiner rather than the noun. It will

[2] For an overview see the entry on *so* in König et al. (1990).

[3] When combined with gradable nouns and verbs, *so* indicates a degree instead of a quality or manner:
(a) *Mein Nachbar ist auch so ein Idiot*. 'My neighbor is an idiot (to a degree) like this guy.'
(b) *Mein Nachbar rast auch so.* 'My neighbor speeds like this, too.'

be shown, however, that the semantic effect is the same regardless of whether *so* is taken to modify the determiner or the noun (cf. section 2.4). The German demonstrative *solch* 'such', which is identical in meaning to *so* but is licensed only with nominal expressions, can in fact precede the determiner as well as the noun without there being a difference in meaning, cf. (3). For ease of exposition the occurrence of *so* in (1b)/(2b) will be considered to be ad-nominal.

(3) a. so ein Auto
b. solch ein Auto
c. ein solches Auto
'such a car / a car like this'

The distributional pattern of the deictic and the anaphoric use of the demonstrative *so* recurs in equative constructions and similes, as shown in (4) – as before *so* combines with adjectives, nouns and verbs. Although we will not go into equatives and similes in this paper, it is worth pointing out that if the standard of comparison is provided by a deictic expression these constructions are equivalent in meaning to the deictic examples in (1). So the variant with a deictic standard of comparison in (4a) is equivalent in meaning to the example in (1a), just as (4b) / (4c) are equivalent (1b) / (1c), respectively. This suggests that the role of *so* in equatives and similes is close to its role in the deictic cases, and that the meaning of *so* is composed of the meaning of *wie* 'like' and deictic reference to an object or event. Assuming that *wie* just expresses similarity, the equivalences between (1) and (4) can be seen as additional support for the analysis suggested in this paper – similarity demonstratives express similarity to the individual or event the speaker points at. (Moreover, the equivalences between (1) and (4) point to a transparent analysis of equatives and similes, which is, however, beyond the scope of this paper.)

(4) <u>equatives</u>
a. Anna ist so groß wie diese Person.
'Anna is as tall as this person.'

b. Anna hat so ein Auto wie dieses Auto.
'Anna has the same car as this car.'

c. Anna hat den Fisch so zerlegt, wie diese Person es tut.
'Anna cut the fish like this person did.'

The main evidence for the claim that the speaker does not point to a kind but instead to an individual or event stems from generics. Compare the demonstrative NP *dieses Auto* 'this car' in (5a) and *dieser Tisch* 'this table' in (6a). While the former allows for a token reading as well as a generic/type reading, the latter only allows for a token reading. This is explained with the help of findings showing that generic readings of definite NPs require well-established kinds (cp. *the Coke bottle* to *the green bottle*, Krifka et al. 1995, p.11). In (5a), a type reading is available regardless of what the context looks like because car subkinds are general knowledge. In (6a), a type reading is not possible in the context of a bar. But it would be possible in other contexts, e.g., shopping at Ikea, where table subkinds are conventional. In contrast, in the case of *so* there is no restriction to well-established kinds: (5b) and (6b) both mean that Anna wants to have a car / table similar to the one the speaker points to.

(5) (speaker pointing to a car in the street):
a. Dieses Auto will Anna haben.
b. So ein Auto will Anna haben.
'Anna wants to have this car / such a car.'

(6) (speaker pointing to a table in a bar):
   a. Diesen Tisch will Anna haben.
   b. So einen Tisch will Anna haben.
      'Anna wants to have this table / such a table.'

The lack of the well-established kind requirement in the case of the demonstrative *so* is evidence that ad-nominal *so*-phrases do not refer to previously established kinds. More evidence that *so*-NPs are not equivalent to *dieser*-NPs on a type reading is provided by the fact that they cannot be subject to a predication by kind-denoting nominals like *Art* 'kind'. While (7a) is unmarked, (7b) is hardly grammatical.

(7) (speaker pointing to a car in the street):
   a. Dieses Auto ist eine besondere Art von Limousine.
   b. ?? So ein Auto ist eine besondere Art von Limousine.
      'This car / such a car is a special kind of limousine.'

There is one more distributional piece of data to be mentioned here: Ad-nominal *so*-phrases cannot occur with definite determiners, cf. (8). This restriction is predicted by the similarity analysis, since *so das Auto* would have to be interpreted as *the unique individual x that is similar with respect to the relevant dimensions to the car pointed at*. Assuming that individuals are similar to themselves (i.e. the similarity relation is reflexive) this is only possible if similarity degenerates to identity, which is excluded by Gricean reasoning, cf. footnote 14.

(8) *so das Auto
    'such the car'

In the current paper the core building blocks of the analysis of similarity demonstratives will be presented focusing on the deictic use of German ad-adjectival and ad-nominal *so*. We will exclude

- first, ad-verbal occurrences of *so*, as in (1c) – the analysis of ad-nominal *so* presumably carries over to the verbal case considering similarity between events instead of individuals. Differences are expected concerning the features qualifying for comparison (cf. section 3);
- secondly, anaphoric uses as in (2) – anaphoric uses differ from deictic uses in that possible antecedents are more diverse than real world referents to be pointed at;
- thirdly, deictic uses based on iconic gestures instead of pointing – as shown in Umbach & Ebert (2009) stretching out one's hands does not represent a measurement in and by itself. For example, the size of a baby will be indicated by a horizontal distance whereas the size of a school kid requires a vertical distance gesture. So even in the case of iconic gestures the demonstration has to account for the spatial orientation of the object (cf. Lang 1989);
- finally, expressions of similarity other than *so* in German and expressions of similarity in other languages will be excluded. English *similar*/*like*/*the same* and German *ähnlich*/*gleich*/*derselbe* differ from each other as well as from the demonstrative *so* in a number of respects, e.g., with respect to which articles are licensed, whether additive particles are licensed, whether there is an NP-internal reading, and whether they are gradable.[4] These issues are beyond the scope of this paper and will

[4]As for additive particles, if the question under discussion is about properties of Anna, say "Anna's weight is 50kg – what is her height?" the speaker may point at a person and answer *so groß ist Anna* / *Anna ist so groß* 'Anna is this tall.' without an additive particle. But if the utterance is out of the blue, the additive particle is required – *so groß ist Anna auch* 'Anna is this tall, too.' which is surprising since there is no antecedent for the additive. Even more

be left for future research (cf. Umbach to appear).[5] It will suffice to include an example showing one striking difference between the demonstrative *so* and the adjective *ähnlich* 'similar'.[6] In (9a) *So ein Geschenk* 'such a present' relates to properties of the present qua being a present, e.g. its size or value. Substituting *ein ähnliches Geschenk* 'a similar present' in (9a) would make the sentences unacceptable, the latter relating to properties of the particular antecedent, e.g., being an exotic animal, cf. (9b). It is thus important to keep in mind that the analysis in this paper addresses the demonstrative *so* but not the adjective *ähnlich* 'similar'.

(9) The prime minister received a Panda bear as a present.
   a. So ein Geschenk /#ein ähnliches Geschenk zeigt die Wertschätzung des Gasts.
      'Such a present / a similar present is evidence of the esteem of the guest.'

   b. Ein ähnliches Geschenk brachte ihm im Vorjahr die Kritik der Tierschützer ein.
      'A similar present evoked protests by animal right activists last year.'

2.2 German *so* and English *so / such* in the literature

While there exists an extensive literature on the non-deictic/non-anaphoric uses of German *so*,[7] the deictic and the anaphoric uses have rarely been addressed. One of the few references to the meaning of the deictic use of *so* is in Ehlich (1987). Ehlich distinguishes five kinds of deixis: speaker/hearer, place, time, objects, and finally what he calls "Aspekte an Objekten" ('aspects of objects'). Following Ehlich the demonstrative *so* is an instance of the 'aspects of objects' kind of deixis and the speaker points to the

---

surprising is the fact that *similar* and *the same* ban additive particles in these contexts – *Anna has the same height, too* cannot be used to convey the information that Anna is of the same height as the person pointed at. As for gradability, in Meier (2009) a comparative semantics for the expression *resemble* is proposed accounting for the fact that it is a gradable predicate (*Mary resembles her mother more than her brother resembles her father*).

[5] There is also a cross-linguistic survey in preparation comparing similarity expressions across a number of languages including Tagalog. Although the results are not yet conclusive, it seems safe to say that Tagalog *gan-ito* ('like this') is a similarity demonstrative occurring in nominal, verbal and adjectival phrases (modulo other forms for distance). When combining with nominals, as in the example below, the linker element *na/-ng* is required.

May ganito-ng kotse rin si Anna.
EXIST like.this-LNK car also SUBJ.PN Anna
'Anna has a car like this, too.'

Following Scontras and Nicolae (in this volume) the linker element is obligatory in cases of non-saturating composition and performs the intersection of the linked arguments, that is, a modification operation. So Tagalog provides independent evidence for the idea of demonstratives acting as modifiers.

[6] One of the reviewers pointed out that German *so* and *ähnlich* can be combined as in *so ein ähnliches Problem* (lit. 'such a similar problem') raising the question of which features are relevant in the combination. This is another question for future research.

[7] To name just a few: In Redder (1987) the correlate function of *so* is analyzed. In Pittner (1993) the role of *so* in insertions is discussed. Hole & Klump (2000) focus on the contraction of *so* and the indefinite article *ein* (*Kaufst du mir so'nen Pullover?* 'Will you buy me a jumper like this?'). In von Heusinger (2012), the contraction of *so* and *ein* is considered as a special indefinite article expressing specificity, and compared to indefinite *this* in English (*Eva will so'n Film über Eliade sehen. / Eva wants to watch this movie about Eliade*.). Wiese et al. (2009) and Jannedy (2010) discuss the distribution of *so* in a particular regional variant of German (*so im Grünen* 'like out in the nature'). Umbach & Ebert (2009) present a semantic analysis of the intensifying use and the hedging use of *so*, which occur "out of the blue", without gesture or antecedent (*Das Hotel ist so teuer.* 'The hotel is so expensive.' / *Der Bischof hatte so'ne Mütze auf.* 'The bishop had like a cap on his head.').

relevant aspects of an object (not to the object itself). In Harweg (1990) the question is raised of whether *so* is a genuine deictic element or rather a cataphor referring to a silent subsequent supplement specifying its meaning. Fricke (2007) comes closest to the position defended in the present paper in classifying *so* as a deictic element similar to object demonstratives like *dieser* 'this', since both require a pointing gesture. She does not explicitly raise the question of the nature of the target of the pointing gesture. In her examples, however, the speaker is reported to point to an object.[8]

In the analysis in König (2012), the field of demonstratives expressing manner, quality, and degree is investigated from a typological perspective. While many languages have demonstratives expressing one (or two) of verbal manner, nominal quality and adjectival degree, demonstratives expressing all of these at the same time (like German *so*, Polish *tak* and Turkish *böyle*) are rare. However, regardless of whether they are as flexible as German *so* or more narrowly confined, like English *such*, all of these expressions function as modifiers of the component they combine with raising the same interpretation problem as the one pointed out for German *so* – how can a demonstrative function as a modifier? They will be subsumed under the notion of *similarity demonstratives.*

While the literature on the deictic and anaphoric use of German *so* is rare, there are some papers on English *so* and *such* to start the semantic analysis from. English *so* and *such* are mostly analogous to German *so* and *solch*. But there are some restrictions. First, English *such*, like German *solch*, can only be combined with nominal expressions. Secondly, English *so*, unlike German *so*, can only be combined with adjectives and verbs. Thirdly, English *so* requires negation or question contexts when occurring in equative comparison (*The old lady is not so innocent as she seems*).[9] One prominent study of English *so* and *such* is Bolinger (1972), distinguishing the identifier use from the intensifier use. The former is deictic or anaphoric (*He is about so tall*. / *We need a telescope equipped for solar photography. Such a telescope is hard to find)* and the latter occurs in result clauses or out of the blue (*He is such a fool that I can't trust him.* / *He was so upset.* / *He was hurrying so.*)

Carlson's (1980) study of reference to kinds includes a semantic analysis of the identifier *such*. Carlson concedes that, at first sight, *such* appears to relate to a modifier antecedent, for example a relative clause (*people who eat fish … such people*). There are, however, a number of problems for a modifier analysis, for example, multiple modifiers (*old ladies who mend shoes … such ladies*), complex noun phrase antecedents (*honest money lenders … such people*), and exclusion of stage-level antecedents (*people in the next room …* ??*such people*) (p. 232). Carlson considers these problems as evidence that *such* relates to kinds instead of modifiers and suggests a semantic interpretation introducing a free kind variable denoting a subkind of the one denoted by the common noun. Siegel (1994) argues, contra Carlson, that *such* is a simple pro-adjective which is not bound by an adjective (or other modifier), but instead by a common noun. This amounts to the claim that although the anaphor *such* functions as a modifier, its antecedent is a nominal phrase. The controversy between Carlson and Siegel was taken up in Landman (2006), who extended Carlson's analysis to the paraphrase *like that* and to ad-nominal German *so* and Polish *tak*. Landman moreover suggests an interpretation of the ad-verbal use of German *so* and Polish *tak* by postulating event-kinds as an ontological category (see also Landman & Morzycki 2003).

In Anderson & Morzycki (2013) this idea is extended by postulating degree-kinds which are thought of as equivalence classes of states holding of individuals of the same degree in some dimension. This allows for interpreting ad-nominal, ad-verbal and even ad-adjectival uses of German *so* and Polish *tak* such that they uniformly refer to kinds – nominal kinds, event kinds and degree kinds, respectively. This analysis is impressively simple but it requires additional ontological stipulations. More importantly,

[8] Here is one of her examples: "*So ein Auto hätte ich gern*. (Zeigegeste auf einen Porsche im Autogeschäft)" 'I want a car like this. (pointing gesture to a Porsche in a car dealer's place.)', Fricke (2007), p. 77.

[9] For a comprehensive overview see Huddleston & Pullum (2002), chap. 13 and 17.

it cannot account for the difference between *so*-phrases and definite generics shown in section 2.1. The problem underlying Anderson & Morzycki's account is their assumption that any collection of all possible objects of some sort is a kind – "The kind RABBIT, for example, consists of all possible rabbits, that is, of all the rabbits in every possible world." (p.12). This is perfect in the case of *rabbit*. But it is problematic in cases where there is no well-established kind, for example in the case of *the green bottle*, cf. section 2.1 and Krifka et al. (1995). In contrast, according to the analysis in the present paper the demonstrative *so* generates similarity classes which are not arbitrary collections of objects. Consider ad-nominal *so*: There is evidence that the 'respects of similarity' have to be features which are criterial with respect to the kind denoted by the noun, cf. section 3.3. Thus the similarity classes generated by ad-nominal *so* are subject to constraints imposed by the concept denoted by the noun. This is good reason to consider the similarity classes as ad-hoc kinds – although they are not 'well-established' kinds, they are not arbitrary collections either, because they have to adhere to requirements imposed by the super-ordinated nominal kind. To avoid misunderstandings, the analysis suggested in this paper does involve kinds. But unlike Anderson & Morzycki's account these kinds need not be given in advance and are instead ad-hoc generated by similarity demonstratives. Thus the analysis in this provides an explanation of how these kinds come into being and how they dovetail with the system of well-established kinds.[10]

2.3 Direct reference – deferred interpretation

Following Kaplan (1989) indexicals, pure indexicals (*I, you, here,* …) as well as demonstratives (*this, that,…*), are directly referential expressions: They receive their interpretation from the context of utterance and have a stable content across world/time indices. The German demonstrative *so* is a directly referential expression in the sense of Kaplan. This is obvious from the fact that counterfactuals shifting the world of evaluation such that the pointing gesture differs from the actual one are not acceptable, cf. (10). In this respect the demonstrative *so* behaves like a run-of-the-mill demonstrative like *dieser* 'this', cf. (11), the unacceptability of shifting the world of evaluation proving that the demonstrative is directly referential.[11]

---

[10] Previous versions of this paper explicitly postulated a sparse ontology including only individuals and events as entities that one can point at. This was meant to support the claim that the target of the demonstration gesture accompanying *so* is an individual or event, but not a kind. It turned out that such an ontology raises issues which this paper is not a place to discuss in a satisfactory way. Since a sparse ontology is not essential for the account pursued in this paper, we dropped it in the final version. So this paper remains as neutral as possible as to what the ontology looks like apart from making a modest suggestion concerning the semantic status of multi-dimensional attribute spaces, cf. 3.3. Still, we are skeptical as to whether the idea of pointing at kinds is reasonable – kinds as well as degrees result from human classification and therefore belong to the realm of mental entities but not to the real world.

[11] Underdetermined pointing gestures can be specified by means of a conditional, as in the example below, which is surprising from a Kaplanian point of view. But since the effect has a meta-linguistic flavor and is parallel for *so* and *dieser* it cannot be taken as an argument against *so* being directly referential. Suppose the speaker cannot properly see which of the two cars she is pointing at:

a. Wenn ich auf das größere Auto zeige, dann hat Anna auch so ein Auto.
   'If I'm pointing to the bigger car, Anna has such a car, too.'
b. Wenn ich auf das größere Auto zeige, dann gehört dieses Auto Anna.
   'If I'm pointing to the bigger car, that's Anna's car.'

Here is an illuminative comment by one of the reviewers: "… what we're seeing here is that Kaplan assumes that certain parameters of the context - who is speaking, what is being pointed at, and so on - are fully determined. But in real life that assumption might not be fully realistic."

(10) (The speaker is facing two cars, pointing to the smaller one)

a. # Wenn ich auf das größere Auto zeigen würde, würde Anna auch so ein Auto besitzen.
'If I were pointing to the bigger car, Anna would own such a car, too.'

b. # Wenn ich auf das größere Auto zeigen würde, wäre Annas Auto auch so groß.
'If I were pointing to the bigger car, Anna's car would be that big.'

(11) (The speaker is facing two cars, pointing to the smaller one)
Wenn ich auf das größere Auto zeigen würde, würde dieses Auto Anna gehören.
'If I were pointing to the bigger car, that would be Anna's car.'

Kaplan's separation of context and index, or character and content of an expression,[12] proved successful and is meanwhile established as the standard interpretation of indexicals. Nunberg (1993), however, pointed out that there are a number of cases that cannot be covered by Kaplan's theory. In (12) the first person pronoun will be interpreted as referring to condemned prisoners in general instead of referring to the speaker only, and in (13) the demonstrative refers to the plates in stock instead of the sample plates the speaker is pointing at.

(12) (Condemned prisoner:)
I am traditionally allowed to order whatever I like for my last meal. (= 32 in Nunberg 1993)

(13) (speaker pointing to sample plates:)
These are over at the warehouse. (adapted from 43 in Nunberg 1993)

For this reason Nunberg dismissed the assumption inherent in Kaplan's system that the target of the pointing gesture must be identical to the interpretation of the demonstrative. The relation between the target of the demonstration and the interpretation of the demonstrative can be some relation other than identity, for example contiguity. Following Nunberg, the semantics of demonstratives as well as pure indexicals includes three components: (i) the target of the demonstration, (ii) the interpretation contributed to the proposition, (iii) some relation between the target of the demonstration and the interpretation of the demonstrative.[13] This accounts for the fact that in (12) the referent of the first person pronoun, i.e. the speaker, is distinct from the interpretation contributed to the proposition, i.e. condemned prisoners in general, and in (13) the target of the pointing gesture, i.e. the sample plates, is distinct from the interpretation of the demonstrative, i.e. the plates in the warehouse.

In Nunberg (1993) cases like (12) and (13) are called "deferred indexical reference". This term is substituted by "descriptive indexicals" in Nunberg (2004), since non-identity between the index (e.g., the speaker in the case of the first person pronoun) and the interpretation is no longer considered as exceptional. While maintaining the earlier analysis as far as demonstratives are concerned – the relation between target and interpretation can be a contiguity relation instead of identity – Nunberg argues for pure indexicals that their descriptive interpretation is not given by contiguity and instead rests on "context granularization" achieved by ignoring a number of features distinguishing the index from the interpretation. This is why, in the case of (12) the first person pronoun can stand for the class of condemned prisoners in general.

[12] The character of an expression is defined as a function from contexts to contents while the content of an expression is a function from indices to denotations.

[13] There is in addition a classificatory component including, for example, number and gender (in the case of pronouns) and a proximal / distal feature (in the case of demonstratives). See also the reformulation of Nunberg's theory in a situation semantic framework by Elbourne (2009).

> "Once we acknowledge that only certain conversationally relevant properties of individuals figure in the domain, it follows that the conversational purposes can determine what counts as an individual, as well. That is, there can be only as many individuals in the domain as are individuated by the conversationally relevant properties – the context is subject to what we can think of as a contextual granularization." (Nunberg 2004, p. 15)

The idea of contextual granularization is strikingly close to the implementation of similarity as indistinguishability with respect to a set of relevant features, which will be suggested for the interpretation of the demonstrative *so* in this paper, cf. section 4.

2.4 The similarity interpretation of the demonstrative *so*

We opted for a solution such that the target of the demonstration gesture accompanying *so* is an individual (or event) – no pointing to kinds and degrees in the case of *so*. On the other hand, similarity demonstratives serve as modifiers, and modification by individuals is impossible (in standard logic). This gave rise to the question we started out from in the introduction: How it is possible that a demonstrative functions as a modifier? The solution is provided by Nunberg's theory of descriptive indexicals, in dismissing the premise that the target of the demonstration gesture and the interpretation contributed to the proposition must be identical (cf. section 2.3). In accordance with Nunberg's theory of descriptive indexicals it will be proposed to interpret similarity demonstratives, and in particular the German demonstrative *so*, by requiring that the relation between the target of the demonstration and the interpretation of the *so*-phrase is similarity instead of identity.

Let us start with ad-nominal cases like (1b), repeated in (14a). It was claimed in section 2.1 that the interpretation of *so* in pre-determiner position yields the same result as the interpretation of *solch* in pre-nominal position. In (14b) the semantics of *so* in a pre-determiner position is shown. The target of the demonstration is represented by a free variable x_{target} to be resolved by the utterance context. Although similarity appears two-place at first sight, it will be argued below that it must be treated as a three-place relation. The third argument, called *F*, represents the features of comparison, that is, the relevant respects of similarity (details will be discussed in section 3). Combining *so* with *ein* 'a' – recall that *so* is restricted to indefinite determiners, cf. section 2.1 – yields the modified determiner in (14c). Adding the noun we get the quantifier in (14d) and finally the sentence meaning in (14e).[14]

(14) a. (speaker pointing to a car in the street):
So ein Auto hat Anna.
'Anna has a car like this.'

b. [[so]] $= \lambda D.\ \lambda P.\ D(\lambda x.\ sim(x, x_{target}, F) \& P(x))$

c. [[so ein]] $= \lambda P.\ \lambda Q.\ \exists x.\ sim(x, x_{target}, F) \& P(x) \& Q(x)$

d. [[so ein Auto]] $= \lambda Q.\ \exists x.\ sim(x, x_{target}, F) \& car(x) \& Q(x)$

e. [[so ein Auto hat Anna]] $= \exists x.\ sim(x, x_{target}, F) \& car(x) \& own(anna, x)$

[14] Intuitively, similarity presumes non-identity of the referent of an ad-nominal *so*-phrase and the target of the demonstration, $x \neq x_{target}$. It can be argued, however, that non-identity is an implicature due to Gricean reasoning. In fact, it seems to be cancelled in:

So ein Auto hat Anna, vielleicht ist das sogar Anna's Auto.
'Anna has a car like this, maybe this is in fact Anna's car.'

Similarity will be implemented as an equivalence relation in section 4, thereby entailing reflexivity. If the non-identity condition is considered as a proper part of the meaning of *so* instead of a mere conversational implicature, it can be added to the definitions in (14) – (16) by constraints , $x \neq x_{target}$.

In (15a) the interpretation of the nominal modifier *solch* is shown. When combining with a noun and an indefinite determiner, the resulting quantifier is identical to the quantifier in the case of *so*, since for an intersectively added restriction it makes no difference whether it is imposed on the domain or on the noun denotation – (14d) and (15c) are identical. This explains why *so ein Auto* is equivalent in meaning to *ein solches Auto*, and it justifies speaking of *ad-nominal so* even if it occurs in the pre-determiner position.

(15) a. [[solch]] = $\lambda P.\ \lambda x.\ sim(x, x_{target}, F) \& P(x)$
b. [[solches Auto]] = $\lambda x.\ sim(x, x_{target}, F) \& car(x)$
c. [[ein solches Auto]] = $\lambda Q.\ \exists x.\ sim(x, x_{target}, F) \& car(x) \& Q(x)$

The ad-adjectival version of the deictic use of *so* differs from the ad-nominal case first of all in the type of the item *so* combines with – in the ad-nominal case it is a determiner or a noun, but in the ad-adjectival case it is an adjective, cf. (16). This results in different types. But beyond, in the ad-nominal case the features (or 'dimensions', or 'respects') of comparison are implicit unless specified in an additional phrase (*with respect to …*) whereas in the ad-adjectival case the feature of comparison is explicit (unless elliptical). In *so ein Auto* there will be several features of comparison and they have to be provided by the context whereas in *so groß* there is only one and it is given by the lexical meaning of the adjective. Thus, while in the ad-nominal case the features of comparison are represented by a free variable *F*, in the ad-adjectival version the feature of comparison is represented by a lambda-bound variable *f* taking adjectival measure functions as arguments (for details about the features of comparison cf. section 3).

(16) a. (speaker pointing to a person):
So groß ist Anna.
'Anna is this tall.'
b. [[so]] = $\lambda f.\ \lambda x.\ sim(x, x_{target}, f)$
c. [[so groß]] = $\lambda x.\ sim(x, x_{target}, height)$
d. [[so groß ist Anna]] = $sim(anna, x_{target}, height)$

Comparing Nunberg's examples of deferred reference to similarly demonstratives like German *so* , there is a crucial difference. Nunberg's examples are systematically ambiguous between an identity interpretation and a descriptive interpretation, even if in some examples the identity interpretation is ruled out by factual knowledge – *these (plates)* in (13) cannot refer to the plates the speaker points to, because plates cannot simultaneously be in front of the speaker and over at the warehouse. In contrast, similarity demonstratives never have an identity interpretation – '*So ein Auto hat Anna.'* cannot have the meaning that Anna owns the car pointed at (unless the implicature of non-identity is cancelled, cf. footnote 14).

Furthermore, according to Nunberg (2004), descriptive demonstratives (but not pure indexicals) are related to their target of demonstration by a contiguity relation, as for example between a car and its wheels, or a book and its publisher. Similarity demonstratives differ from Nunberg's descriptive demonstratives since similarity is clearly no contiguity relation – there is no similarity between, e.g., a car and its wheels. What is more, similarity demonstratives can themselves rely on contiguity. In (17) the demonstrative *so* imposes similarity instead of identity. At the same time the target of the demonstration is a book but the noun denotes a publishing house. Thus in (17) similarity and contiguity

are combined – the referent of the *so* NP has to be a publishing house similar to the publishing house of the book the speaker points at.

(17) a. (speaker pointing to a book)
So ein Verlag ist selten.
'Such a publishing house is rarely found.'

b. [[so ein Verlag]] = $\lambda Q.\ \exists x.\ sim(x, y, F)\ \&\ contiguous(y, x_{target})\ \&\ publisher(y)\ \&\ Q(x)$

According to Nunberg, descriptive demonstratives are interpreted by contiguity whereas descriptive (pure) indexicals are interpreted by *contextual granularization* conflating indistinguishable entities (cf. section 2.2). *Contextual granularization* comes very close to replacing identity by similarity – similarity will in fact be spelled out by indistinguishability in section 4. From this perspective, similarity demonstratives pattern with descriptive indexicals instead of (regular) demonstratives.[15] But unlike indexicals they may in addition include contiguity, as in (17). This is plausible taking into account that contiguity is a matter of the target of the pointing gesture (book or publishing house) while contextual granularization is a matter of how the referent relates to a given target of pointing (identity or similarity).

The interpretation of similarity demonstratives via similarity between the target of the demonstration and the interpretation of the *so*-phrase is a simple and, from the point of view of ontology, conservative solution. But the similarity relation in itself is not unproblematic. Goodman (1972) decidedly argued against similarity: "Similarity, ever ready to solve philosophical problems and overcome obstacles, is a pretender, an impostor, a quack." (p. 437). One of his objections is 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.). Goodman concludes "We have to say […] in what respects two things are similar." (444). But then he seems to confuse respects and properties: "To say that two things are similar in having a specified property in common is to say nothing more than that they have that property in common." (445). In our view, respects (or 'dimensions' or 'features') constitute a property only if combined with a value. For example, the respect *color* turns into a (first-order) property only if combined with a value like *red, green*, *blue*, etc. Contra Goodman: To say that two things are similar with respect to a certain respect, say *color*, is not equivalent to saying that they share a property, say *being red* (or *green*, *blue*, etc.). Dimensions/features/respects – we use these terms interchangeably – require a value to form a property (*be red* …). This is the view employed in the following sections.

3 Multi-dimensional comparison

3.1 Adjectives vs. nouns

The notion of similarity is trivial if not relativized to particular respects of similarity, and a two-place similarity relation would be insufficient – this is the lesson from Goodman's objections against similarity. But how can we find out which respects are relevant? In the case of ad-adjectival *so* the answer is obvious: The dimension of comparison is given by the adjective's meaning – *tall* relates to the dimension of height just as *heavy* relates to the dimension of weight and *old* relates to the dimension of age. There is, however, no one-to-one-correspondence: Some dimensions are targeted by more than one adjective,

[15] In fact, in German a sentence like *'I am traditionally allowed …'* would be translated by means of *so*: '*So einer wie ich* …' ('such a guy as I am …').

for example *wide*, *deep*, and *high* (as applied to a cupboard) likewise relate to the dimension of (physical) length. At the same time, some adjectives relate to more than one dimension. For example, *large* applied to a city may relate to the area or to the number of inhabitants. Note, however, that in an actual comparison these two dimensions cannot co-occur – if you have to decide whether Berlin is larger than Hamburg, you have to decide on either the area or the number of inhabitants. This is clear evidence that adjectives relating to more than one dimension are ambiguous, the dimensions specifying different readings.[16] Thus for adjectives there is only one dimension involved in an actual comparison.[17] Unlike adjectives, in the case of nouns a single comparison may target more than one dimension. Similarity of Anna's car to the one the speaker points to may simultaneously relate to the dimensions of *drive type*, *horsepower*, *number of doors*, *technical equipment* etc. Adjectival dimensions, at least those targeted by adjectives like *tall* and *heavy*, relate to metrical scales. Unlike adjectival dimensions, nominal dimensions need not be metrical. They may relate to scales of lower levels of measurement, viz. ordinal (the values being non-metric but ordered) or nominal (the values being elements of an unordered set, for example, drive types including *diesel*, *gasoline*, *natural gas*, and *electric*). Nominal scales may even consist of only two values, that is, be binary. [18]

The difference between adjectives and nouns – one-dimensionality vs. multi-dimensionality – is demonstrated in (18) and (19). In the ad-nominal example in (18a) the second speaker's request for clarification is unmarked since there are various dimensions to choose from, whereas in the ad-adjectival example in (18b) an analogous clarification request is infelicitous because the one dimension of comparison is already given by the adjective (B's request would only be felicitous when referring to different readings of *groß*, e.g., *tall* vs. *great*). Likewise, in the nominal version in (19a) the causal clause contains more than one conjunct in reasoning why Anna's car is like the car the speaker points at. In the adjectival version in (19b), an analogous construction is infelicitous.

(18) a. A: (pointing to a car in the street):
So ein Auto ist Annas auch.
'Anna's car is one like this.'
B: In welcher Hinsicht?
'In which respect?'

b. A: (pointing to a person):
So groß ist Anna auch.
'Anna is this tall, too.'
B: # In welcher Hinsicht?
'In which respect?'

[16] Sassoon (2011) argues that adjectives may have more than one dimension, for example, *healthy with respect to blood pressure, cholesterol, sugar* etc. She does not claim, however, that there may be more than one dimension involved in an actual comparison. On the contrary, she argues that adjectival dimensions are integrated by logical operations while nominal dimensions are integrated through similarity. Integrating adjectival dimensions by, e.g., conjunction, amounts to considering them one-by-one, yielding a list of (simple) comparisons: Berlin is larger than Hamburg with respect to the number of inhabitants but not with respect to the area. From this point of view, her distinction between adjectives and nouns is close to the one proposed in this paper.

[17] According to Gärdenfors (2000) adjectives may relate to more than one dimension constituting an integrated sub-domain. Color adjectives may, for example, be viewed as having three sub-dimensions (e.g. in the RGB model). This is evident from the point of view of physics, but when considering linguistic expressions specifying dimensions there is no way of distinguishing sub-dimensions – it is nonsensical to say that something is pink with respect to its blue-component but not with respect to its red-component.

[18] Levels of measurement are basic in statistics, details can be found in the handbooks.

(19) a. Annas Auto ist so wie das da, weil es auch ein 5-Türer ist und Heckflossen hat.
'Anna's car is like this one because it has also 5 doors and tail fins.'

b. ?? Anna ist so groß wie diese Person, weil sie auch 1.80 ist und 5 Grammys gewonnen hat.
'Anna is as tall/great as this person because she is also 180 and won 5 Grammys.'

The idea of multi-dimensionality of comparison is fundamental for the analysis of so-called similarity comparatives in Alrenga (2007). Alrenga investigates constructions headed by *different*, *same* and *like*, which are in many respects parallel to adjectival comparatives headed by *more* /*-er* and *as*. But while the latter relate to only one dimension, the former relate to more than one dimension simultaneously. This is shown in (20a-c) (=61 in Alrenga 2007) where dimensions are explicitly named.

(20) a. In all relevant respects, this place is the same as it was before.
b. My new car is a lot different than my old one was: it looks different, it drives differently, and it costs more money.
c. Apples are like oranges in that they are round, edible, have seeds, etc.

The interpretation of similarity comparatives proposed by Alrenga is analogous to the degree-based account of adjectival comparatives (e.g. Kennedy 1999): Similarity comparatives measure the degree of (dis)similarity between two individuals based upon the differences amongst the individuals' values along the dimensions of comparison. The denotation of *different* is stated in terms of a measure function μ_{DIS} mapping individuals to degrees. But unlike scalar comparison these degrees are not degrees of a concrete dimension, like HEIGHT or WEIGHT, and are instead degrees of a fixed relation R merging all contextual relevant dimensions into one dimension of DISSIMILARITY. For example, the sentence *Barry is different than I am* is interpreted analogous to the sentence *Barry is taller than I am*. But while the latter is about a difference in height, the former is about a difference with respect to a number of contextual relevant dimensions merged in the constant relation R.

Alrenga's account is half-way towards the generalized measure function account aimed at in this paper. However, the combination of dimensions and thus the nature of the similarity measure represented by the DISSIMILARITY dimension is a black box in Alrenga's account. Although he briefly refers to feature matching approaches, e.g. Tversky 1977, the (dis-)similarity relation is treated as a semantic primitive, thereby precluding the possibility of distinguishing between different respects in which individuals are same or different, and ruling out inferences such as "*A is taller than B. Therefore A is different from B with respect to height*.", which is a serious drawback.

3.2 Constraints on dimensions

As argued in the previous section, similarity involves one dimension in adjectival comparison and many dimensions in nominal comparison. In the adjectival case the dimension is easily retrieved from the lexical meaning of the adjective. But what about the nominal case? How can the relevant car-dimensions in (14) be determined? The first and easiest answer would be to refer to the context. But fortunately, there seem to be constraints given by the lexical meaning of the noun. A's reply to B in (21a)

is unmarked because having a gas engine as well as having a hatch is essential for cars. (21b) is still unmarked because being dented is a typical appearance of cars. (21c), however, appears marked because having a CD-player is not essential for cars.

(21) 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 und eine Ladeklappe.
'Anna's car also has a gas drive and a hatch.'

b. A': Anna's Auto ist auch vollkommen verbeult.
'Anna's car is also heavily dented.'

c. A': ??? Anna's Auto hat auch einen CD-Spieler.
'Anna's car also has a CD player.'

Strong contextual support can make a dimension relevant for comparison. In the context in (22) in-car entertainment is highly relevant making the CD-player available for comparison. This is why B's answer appears acceptable.[19] In the example in (23) the relevant property of the car A points at is that it got a parking ticket. Unlike the CD case, where the relevant dimension is *in-car entertainment* there is not even a name for such a dimension (*stuff on the windshield* ?) and in fact B's anwer in (23) is marked in spite of strong contextual support.[20]

(22) A: Ich will in den Ferien über Land fahren und an die alten Zeiten denken und dabei meine ganzen alten Kassetten hören. Es gibt aber ein Problem: Mein Auto hat nur einen CD-Spieler. Wer kann mir helfen?
' I'm planning a retro road trip, complete with my collection of 8-track tapes. Only one problem: this car only has a CD player. Who can help me out?'
B: Ich leider nicht. Ich hab auch so ein Auto.
'Not me; I have such a car, too.'

(23) 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.'

The effect in (21) – (23) is reminiscient of the examples of the infelicitous use of *such* discussed in Carlson (1980), e.g. *people in the next room* … ?? *Such people …*, cf. section 2.2, where the attribute *in the next room* is said to be unfit for selecting a subkind. This raises the question of how to separate kind-selecting attributes from others, which is taken up in Carlson (2010) exploiting the relation between

[19] First, we would like to thank one of the reviewers for this example. Secondly, our informants were mixed: Some found (22) completely unmarked and some said that it requires an extra accommodation effort.

[20] It has to be noted that in (22) and (23) the demonstrative is used anaphorically. Anaphoric uses might be more liberal with respect to dimensions of comparison.

generic sentences, kinds, and (psychological) concepts. Carlson refers to an experimental study by Prasada and Dillingham (2006) who found that humans represent principled connections between concepts that correspond to kinds and some, but not all, of the concept's properties. Prasada and Dillingham distinguish *k-properties* from *t-properties*, the former being properties humans ascribe to entities because they are the kind of things they are, and the latter including factual and statistical properties.[21] The difference between the two types of properties is demonstrated in (24) and (25): (24a) can be paraphrased by (b) as well as (c), whereas (25a) can only be paraphrased by (b) the paraphrase in (25c) being unacceptable. Although most barns are red (according to Prasada and Dillingham), *being red* is not a property of barns because they are barns. In contrast, *being four-legged* is a property of dogs because they are what they are, even if there are some three-legged dogs (cf. (1) and (2) in Prasada and Dillingham 2006).

(24) a. Dogs are four-legged.
b. Dogs, in general, are four-legged.
c. Dogs, by virtue of being the kinds of things they are, are four-legged.

(25) a. Barns are red.
b. Barns, in general, are red.
c. # Barns, by virtue of being the kinds of things they are, are red.

In their study, Prasada and Dillingham test possible paraphrases (as in 24/25), possible explanations for why an entity has a given property (*Why does that* (pointing to a dog) *have four legs*?), and also normative statements (*Dogs should have four legs*.). Their results unambiguously confirm the distinction between k-properties and t-properties. This is evidence that concepts / kinds denoted by (simple or compound) nouns are connected to certain properties in a principled way, independent of frequency effects. In addition, Prasada and Dillingham argue that these properties are directly connected to the respective kind, that is, if a property is a k-property of a given kind, it is not a k-property of the superordinate kind. Put it the other way around, if a property is inherited from a superordinate kind, it is not a k-property anymore. Evidence is provided by examples of the form of (26).

(26) a. #Dogs, by virtue of being dogs, are extended in three dimensions.
b. Dogs, by virtue of being material beings, are extended in three dimensions.

From a linguistic point of view, the distinction between k-properties and t-properties yields an explanation of why bare plurals in generic sentences are not always substitutable by indefinite singular versions (# *A barn is red.*). (See Carlson 2010, who links k-properties to Greenberg's 2003 notion of "principled connection", and see also Krifka 2012, who exploits the normative character of k-properties interpreting indefinite singular generic sentences as definitions.)

Coming back to ad-nominal *so* and the observation that having a gas engine is good and having a CD player is not always good and having a parking ticket is bad as a dimension of comparison, Prasada and Dillingham's findings provide a plausible explanation: The first is a k-property while the third is not and

[21] The k in k-properties stands for *kind*, the t in t-properties stands for *type*, i.e. arbitrary classifications.

the second requires accommodation as a k-property. Unfortunately, we cannot straightforwardly adopt Prasada & Dillingham's k-properties as dimensions of comparison because dimensions must not include specific values. Recall that properties are formed by dimension plus value: *Having four legs* is a property while *number of legs* is the corresponding dimension.

Fortunately, Prasada & Dillingham's "by-virtue-of-the-kind-it-is" paraphrase can easily be adapted to dimensions. What we need are dimensions connected to a kind in a principled way. This suggests to select dimensions ascribed to entities by virtue of the kind they are. Dimensions like *number of doors* or *type of drive* are perfect. A dimension like *entertainment devices* requires a particular context to qualify for comparison, and *stuff on the windshield* is unsuited because cars do not have stuff on the windshield in virtue of being cars. So it seems that the dimensions required for comparison in ad-nominal *so*-phrases are dimensions which yield a k-property for the particular subset denoted by the *so*-phrase. Such dimensions will be called *criterial dimensions* in this paper.

When interpreting ad-nominal *so*-phrases the dimensions of comparison will be restricted to the criterial dimensions of the kind denoted by the noun. This restriction entails that the set of elements in a similarity class generated by *so* is not just an arbitrary subset of the noun denotation. It is instead characterized by means of k-properties and thus eligible to be called a subkind. But it need not be a previously established subkind in a well-established taxonomy (cf. the examples in (5)/(6)) and it need not be a subkind denoted by a noun or compound – there is no lexicalized word like *Gasautos* (lit: 'gas car') in German. The subkind denoted by a *so*-phrase is ad-hoc generated by similarity to the target of the demonstration. This leads us back to the subkind interpretation of English *such* (Carlson 1980) and German *so* / Polish *tak* (Anderson & Morzycki 2013). The analysis in this paper shows that ad-nominal *so*-phrases do in fact denote subkinds. However, interpreting *so* by similarity and considering criterial dimensions reveals why *so*-phrases do that. It has been objected that the question of how similarity relations create categories can be left to the categorization theorists while the semanticist can simply proceed by invoking kinds in their analysis. But then, demonstratives like *so/such/tak* would lack a transparent analysis.

3.3 Generalized measure functions

Summarizing the findings up to now, it was argued in section 2 that the demonstrative *so* in its deictic use is directly referential but the relation between the target of the demonstration and the referent of the *so*-phrase is not identity but instead similarity. We focused on ad-adjectival and on ad-nominal occurrences of *so* (cf. 14, 16 in section 2.4) setting ad-verbal occurrences aside. In the ad-nominal case – *so ein Auto* 'such a car' – the target of the demonstration is the car the speaker points at, and the referent of the *so*-phrase is a car similar to the target. In the ad-adjectival case – *so groß* 'that tall' – the target of the demonstration is the person the speaker points to and the referent that the *so*-phrase is predicated of is similar to the target with respect to height. It would be unsatisfactory, however, if similarity is introduced as a semantic primitive or an artificial dimension of (dis)similarity as in Alrenga (2007). The approach presented in this paper aims at a transparent notion of similarity based on multi-dimensional spaces. A well-known instance of multi-dimensional spaces are conceptual spaces as suggested in Gärdenfors (2000). Unlike Gärdenfors' conceptual spaces, which employ a quantitative similarity measure (geometrical distance), we will employ a qualitative notion of similarity understood as indistinguishability with respect to properties defined on dimensions (which is close to Nunberg's idea

of 'contextual granularization', cf. section 4).[22] The multi-dimensional spaces this measure is implemented on will be called *attribute spaces* in section 4. Another fundamental difference between attribute spaces and Gärdenfors' conceptual spaces is their status: Gärdenfors' conceptual spaces form a stand-alone system, independent of truth-conditional semantics. In contrast, we want attribute spaces to be integrated into truth-conditional semantics and we want a notion of similarity suited for compositional interpretation.

Integration is simple: In their standard degree-based interpretation gradable adjectives are associated with measure functions from (or relations between) individuals to degrees on a scale representing the adjective's dimension. The adjective *tall*, for example, is associated with a function *height* mapping individuals to degrees of height. Scales used by adjectival dimensions are metrical, i.e. the values can be measured by real or natural numbers.[23] Unlike adjectives, nouns are multi-dimensional and the dimensions need not have metrical scales – scales may also be ordinal or nominal or even binary (see section 3.1). From this point of view, nominal and adjectival comparison differ only with respect to the number of dimensions and the nature of scales.

This suggests generalizing the well-established notion of adjectival measure functions. Adjectival measure functions are one-dimensional, mapping individuals to degrees, i.e. values in the adjectival dimension. Generalization to more than one dimension yields functions mapping individuals to values in each of the dimensions, that is, mapping individuals to points in multi-dimensional spaces. These functions will be termed *generalized measure functions*.

(27) and (28) show a one-dimensional measure function and a multi-dimensional one. The adjective *tall* in (27) is associated with the dimension of HEIGHT which is measured by real numbers. The noun *car* is (in this example) associated with the dimensions TYPE OF DRIVE, NUMBER OF DOORS, TECHNICAL EQUIPMENT, HORSEPOWER, and ELECTRONIC IMMOBILIZER, which are supposed to be criterial dimensions of the kind 'car'. The dimension of horsepower and the number of doors have metrical scales, but the drive types are nominal values, the scale of technical equipment is ordered along the partial order of subsets, and the electronic immobilizer dimension is binary.

(27) One-dimensional measure function associated with *tall*:

μ_{height}: $U \rightarrow \mathfrak{R}$

(28) Many-dimensional measure function associated with *car*:

DRIVE_TYPE:	$U \rightarrow$ {diesel, gasoline, natural gas, electric}
NUMBER OF DOORS:	$U \rightarrow \{1 \ldots 5\}$
EQUIPMENT:	$U \rightarrow \wp$ {rear assistance, lane guide, park pilot, BLIS}
HORSEPOWER:	$U \rightarrow \mathfrak{R}^+$
ELECTRONIC IMMOBILIZER:	$U \rightarrow \{0, 1\}$

Instead of writing it as a feature structure, as in (28), the multi-dimensional measure function may also be written as a function into tuples of values, as in (29). Note that while in the one-dimensional case

[22] In using a qualitative notion of similarity based on features/dimensions instead of distance this approach is close to Tversky's (1977) contrast model of similarity. Still, the notion of similarity in the contrast model differs substantially from the one defined in section 4.3. In the contrast model, similarity is a linear combination of the number (or the measures) of the common and distinct features such that similarity increases with addition of common features and/or deletion of distinctive features. This yields a gradable notion of similarity, which is correct since Tversky meant to capture the meaning of the adjective *similar*. *So*/*solch* is not gradable and differs from *ähnlich/ similar* in a number of other respects (cf. section 2.1 and 4.3).

[23] Unless they are evaluative, like *beautiful* and *tasty*.

measure function and dimension are identical (μ_{height} in (27) is in fact the dimension of height), in the multi-dimensional case a measure function is composed of a number of basic functions, cf. (28).

(29) μ_{car}: U $\rightarrow$ <x_1, x_2, x_3, x_4, x_5>,
where $x_1 \in$ { diesel, …}, $x_2 \in \{1…5\}$, $x_3 \in \wp$ {rear assistance,…}, $x_4 \in \Re^+$, $x_5 \in \{0,1\}$

From a technical point of view, generalized measure functions are innocent. They can be viewed as a simple version of feature structures (as, e.g., used in HPSG). What is critical is their status within semantic interpretation. The analysis in this paper is first of all based on standard truth-conditional semantics with a realistic notion of reference. This suggests that feature values, numerical values as well as nominal ones, are not part of the domain. But then, what about their semantic status?

A satisfactory answer to this question would be beyond the scope of this paper – all we can do here is take the discussion about degrees as a guideline. Since multi-dimensional measure functions are a generalization of degree functions, stipulations about degrees carry over to other values. Most semantic accounts of gradability making use of degrees view them as "abstract representations of measurement" (e.g., Kennedy 1999, p. 49). Cresswell (1976) suggested to view degrees as representations of equivalence classes, but even then they are abstract entities, ontologically distinct from individuals. In fact, degrees are generally considered to be of a distinct semantic type (type d instead of e).[24]

Bierwisch (1987) is one of the rare authors touching upon the status of degrees. He considers degrees as being generated by comparison:

> "Vergleichsoperation und Grade einer Skala bedingen sich gegenseitig: Ohne Vergleich keine Grade, ohne Grade kein Vergleich."[25] (p. 130)

Following Bierwisch, degrees are mental entities. This entails that they are not part of a realistic ontology, and in particular, pointing to degrees is impossible. They are auxiliary entities required, e.g., by comparison, and are inherently connected to their dimension – degrees can in fact be viewed as pairs of dimension and value. This perspective seems reasonable for other dimensions, too. Nominal dimensions, for example DRIVE_TYPE, are best viewed as mental entities, auxiliary in specifying car subkinds. Thus, although the interpretation in this paper is grounded in standard truth-conditional semantics, we suggest to borrow from cognitive semantics degrees and other dimensional values as mental entities invoked ad-hoc by comparison. Since these entities belong to the realm of concepts or senses (cf. Carlson 2010) the question arises of how they can figure in the denotations of predicates constructed out of individuals. We do not have a conclusive answer. Still, the setting introduced in section 4 can be considered as a first try. It includes classification functions defined on attribute spaces mirroring predicates on individuals. For example, there is a classification function *tall** corresponding to the predicate *tall* such that *tall*(μ_{height}(x))* is true (approximately[26]) if and only if *tall(x)*. This is why the diagram in figure 1 has to (approximately) commute. Classification functions can be seen as linking mental attribute spaces to real world denotations where attribute spaces provide an auxiliary means in determining the truth of a predicate.

[24] Moltmann (2005) refers to tropes instead of degrees while considering tropes to be entities of a realistic ontology. However, pointing to tropes does not appear more conclusive from a cognitive point of view than pointing to degrees.

[25] "The operation of comparison and the degrees of a scale are mutually dependent: No degrees without comparison, no comparison without degrees." – this quote has to be read in the context of gradability of adjectives; Bierwisch did not consider similes in his paper.

[26] Since classification functions are fuzzy there is no exact equivalence.

4 Similarity in multi-dimensional attribute spaces

4.1 Multi-dimensional attribute spaces

In section 2.4 interpretations of ad-nominal *so* and of ad-adjectival *so* were proposed such that the demonstrative denotes a similarity relation, cf. (14b), (16b), repeated in (30). In this section, the similarity relation will be spelled out in a setting including multi-dimensional attribute spaces. (For ease of exposition, the interpretation of adjectival *so* in (30a) differs slightly from that given in (16) representing the dimension of comparison as a singleton set of dimensions F, i.e. F = {f}, where F is a free variable instead of the lambda bound one in (16)).

(30) a. ad-adjectival *so*
[[so]] = $\lambda x.\ sim(x, x_{target}, F)$

b. ad-nominal (ad-determiner) *so*
[[so]] = $\lambda D.\ \lambda P.\ D(\lambda x.\ sim(x, x_{target}, F)\ \&\ P(x))$

The setting proposed in this paper is shown in figure 1. It includes, (i) a domain ***D*** which is a subset of the universe; (ii) generalized measure functions μ defined on ***D***; (iii) attribute spaces ***F*** spanned by the relevant dimensions; (iv) classification functions p^*_D corresponding to natural language predicates. A measure function maps elements of the domain to points in the relevant attribute space. In the case of adjectives the attribute space is one-dimensional.[27] The classification functions map points in the attribute space onto truth values. Note that dimensions are associated with classification functions (see 4.2 below). The set of possible classification functions is considered as a component of an attribute space ***F*** and is denoted by ***C*(*F*)**. Functions p_D represent regular predicates taking individuals and returning truth values. Measure functions and classification functions have to be such that the diagram (approximately) commutes, that is, $p_D(x) = p^*_D(\mu(x))$.

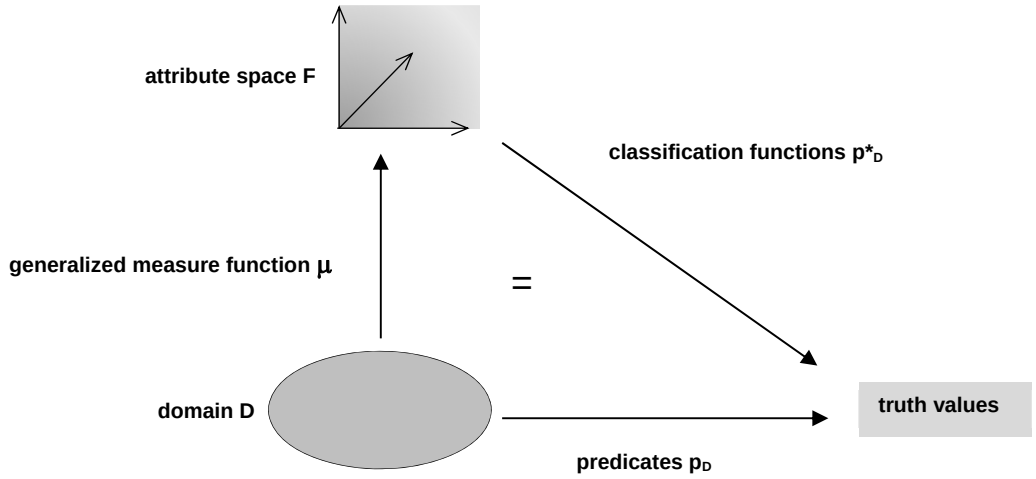

Figure 1: The combination of domain and attribute space

[27] According to Gärdenfors (2000) adjectives may relate to more than one dimension constituting, however, a natural property forming a sub-domain. Color adjectives may, for example, be viewed as having three sub-dimensions (e.q. in the RGB model). But this is a point of view of physics, not of linguistics. See also section 3.1.

Consider the one-dimensional adjectival case, for example the predicate *tall* applied to human adults. The measure function μ_{height} maps individuals in the domain of human adults to degrees of height, i.e. real numbers. The classification function $tall^*_{human\_adults}$ implements the cut-off point determining which degrees of height count as tall in the domain of human adults. Since the diagram (approximately) commutes, $\lambda x.\ tall^*_{human\_adults}(\mu_{height}(x))$ (approximately) implements the meaning of *tall for human adults*.

Generalization to the nominal case is straightforward. Suppose there is a set of dimensions constituting the attribute space (cf. section 3.2). In the nominal case dimensions may relate to all sorts of scales – metrical as well as ordinal and nominal ones. Measure functions will map the individuals in the domain under debate to points in the attribute space (cf. section 3.3). As before, attribute spaces are equipped with classification functions which may combine into complex classifications, see the example of *high-powered diesel car* in (34) below.

4.2 Classification functions

In the setting described above the cut-off for the positive form of, e.g., *tall* is not given as a degree but instead by a classification function $tall^* \in C(F_{height})$. Thus Anna counts as tall iff $tall^*(\mu_{height}(anna))$ is true. Similarly, a measure phrase like *1.80m* will not be interpreted as a point on the scale of height but instead as a classification function $1.80m^* \in C(F_{height})$. The reason is that the proposition in (31), for example, does not mean that Anna's height is exactly 1.80000000 …. meter. It may be true even if she is 1.79 or 1.81. Measure phrases are well-known to have a context-dependent granularity or tolerance range (cf. Krifka 2009).

(31) Anna is 1.80 tall.

If classification functions are assumed to be crisp functions, a certain range of tolerance has to be allowed, say +/- 1cm in the case of the 1.80* classification. It seems more natural, though, to model classification functions as fuzzy membership functions (Zadeh 1965). Using fuzzy membership accounts for the vagueness of the predicate *tall* as well as the tolerance range of the measure phrase *1.80.*

Regardless of whether crisp or fuzzy, classification functions are defined by a set of basic membership functions, for example the basic functions *short**, *medium**, *tall**. New classification functions can be constructed from basic ones by a limited set of operations: conjunction, disjunction, negation and closures, where the set of basic membership functions determines the maximal granularity of classification. The three basic functions above yield a rather coarse granularity which may nevertheless be sufficient in many cases. Classification functions should be cognitively plausible and therefore be subject to the constraint that they do not contain 'holes'. This means, "*short** *or tall**" should be a tautology (corresponding to the constant membership function giving 1 or *true*). This constraint corresponds to convexity in geometric spaces and can be modeled in attribute spaces by means of a closure operator.[28] If the points of the attribute space provide a (partially) ordered set, closure operators can be defined in a straightforward way. For a subset *A* (of the points) of an attribute space with partial order we get the definition in (32), which can easily be extended to fuzzy functions. Figure 2 shows fuzzy membership functions *short**, *medium**, *tall**, closure of "*short** *or tall**" and left- as well as right-closure of *medium**.

[28] A closure operator performs completion of sets in some respect. Axioms for a closure operator are $cl(\{\}) = \{\}$, $X \subseteq cl(X)$, $X \subseteq Y \rightarrow cl(X) \subseteq cl(Y)$ and $cl(X) = cl(cl(X))$. Convex closure operators must additionally have the anti-exchange property: If neither y nor z belong to cl(X), but z belongs to $cl(S \cup \{y\})$, then y does not belong to $cl(S \cup \{z\})$.

(32)

$cl(A)$	$= \{x \mid \exists y,z \in A: y \leq x \leq z\}$	closure
$cl_{\leftarrow}(A)$	$= \{x \mid \exists z \in A: x \leq z\}$	left closure
$cl_{\rightarrow}(A)$	$= \{x \mid \exists y \in A: y \leq x\}$	right closure

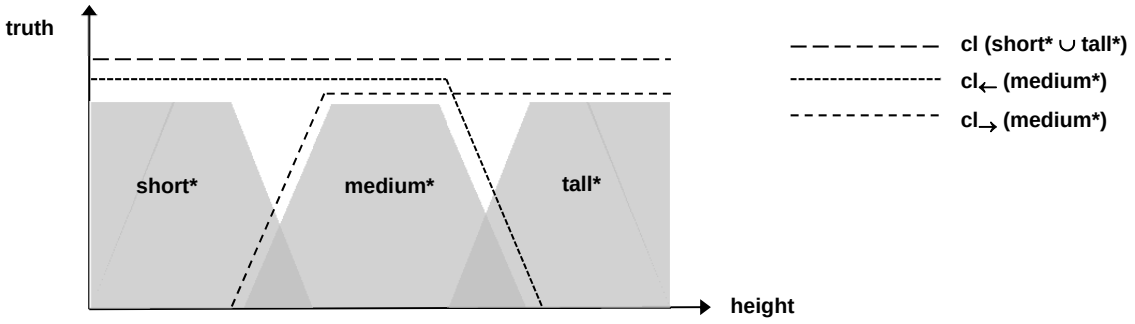

Figure 2: Fuzzy membership as classification function

Summarizing, a dimension is given by a function from a domain into a (partially ordered) set M – possible values of the dimension – together with a set of basic (fuzzy) classification functions and a closure operator. With these prerequisites we can generate a system of classification functions $C(F)$ associated with an attribute space F:[29]

(33) All basic classification functions are in $C(F)$

If $A \in C(F)$ then

$cl(M\text{-}A) \in C(F)$	closure of complement
$cl_{\leftarrow}(A) \in C(F)$	left closure (if partially ordered)
$cl_{\rightarrow}(A) \in C(F)$	right closure (if partially ordered)

If $A,B \in C(F)$ then

$A \cap B \in C(F)$	intersection
$cl(A \cup B) \in C(F)$	closure of union

Here is an example: Let us assume that *horsepower* and *drive_type* are the relevant dimensions. *Horsepower* has a metrical scale isomorphic to $\Re^+$, and *drive_type* has a nominal scale with values {*diesel*, *gasoline*, *natural gas*, *electric*}. We will specify one basic classification function *medium-powered** which is true at least between 70 hp and 90 hp and false at least below 60 hp and above 110 hp. Other classification functions for the *horsepower* dimension will be generated by logical operators and the closure operators coming from the intrinsic order of $\Re^+$, for example *low-powered** as $cl(not(cl_{\rightarrow}(medium\text{-}powered^*)))$ and *high-powered** as $cl(not(cl_{\leftarrow}(medium\text{-}powered^*)))$. For the nominal scale *drive_type*, each value of the scale yields a basic classification function *v**. For example, *diesel** is a classification function which is true on the value *diesel* and false otherwise. For nominal scales closures do not have an effect. The simplest way of constructing cross-dimensional classification functions in $C(F_{car})$ is by pure combinatorics: *high-powered_diesel**(x) = *high-powered**(x) & *diesel**(x). The predicate *high-powered diesel car* can then be approximated by the classification shown below.[30]

[29] Recall that classification functions map elements of M to truth-values and are thus equivalent to subsets of M.

[30] It could be objected that there are dependencies between the dimensions: what counts as high-powered may depend on the drive type. A 90 hp electric drive may count already as high-powered. This means that the classification functions for the *horsepower* dimension may be dependent on the drive type, e.g. by a drive type dependent factor c: *drive_type* $\rightarrow \Re^+$ with c(*gasoline*) = 1.0, c(*natural gas*) = 1.0, c(*diesel*) = 0.8, c(*electric*) = 0.5. We then get a predicate 'high-powered electric car' such that

(34) *high-powered diesel car* (*x*) = *high-powered**($\mu_{\text{horsepower}}(x)$) & *diesel**($\mu_{\text{drive\_type}}$(x))
where $\mu_{\text{Fcar}}(x) = (\mu_{\text{horsepower}}(x), \mu_{\text{drive\_type}}(x))$

Attribute spaces comprising a system of classification functions *C*(*F*) as defined above will be used to model similarity in the next section.

4.3 Similarity as indiscernibility

Standard approaches in Artificial Intelligence employ a similarity measure assigning to each pair of entities a real number between zero and one*, s: D*x*D* $\rightarrow$ [0,1]. There are, however, two problems with such a similarity measure. The first problem is that the elements of a domain *D* do not in general form a nice mathematical space – for example, similarity of cars cannot be accessed in a direct way – and thus it is unclear how to define a similarity measure directly on a domain *D*. In contrast, in our setting a similarity measure can be defined on attribute spaces and transferred to the domain by combining it with the measure function. Suppose $\boldsymbol{s_F}$ is a similarity measure on an attribute space and $\boldsymbol{\mu_F}$ is an appropriate measure function, then a similarity measure $\boldsymbol{s_D}$ on the domain *D* is defined by $s_D(x,y) = s_F(\mu_F(x), \mu_F(y))$.

The second problem is that we do not want a similarity measure returning real numbers. We want instead a similarity predicate returning true or false matching the predicate *sim*(*x*, *y*, *F*) in the interpretation of the demonstrative *so* in (30). The naïve way to turn a similarity measure into a predicate would be by introducing a threshold δ and define the predicate as in (35):

(35) $\text{sim}(x, y, F) = s_F(\mu_F(x), \mu_F(y)) \geq \delta$.

However, introducing a threshold does not give the intended results. First, such a threshold is an arbitrary number not motivated by any interpretational aspects. It is not linked anymore to the criteria that led to similarity and thus makes the wrong predictions. Consider someone saying *So ein Auto hat Anna und so ein Auto hat auch Berta* 'Anna has such a car and Berta has such a car, too.' where the speaker points at the same car throughout. This utterance cannot be interpreted as *Anna has a gas-powered car like this one, and Berta has a four-door car like this one.* But if similarity were measured relative to a numerical threshold such a reading would be predicted to be possible.[31]

Moreover, the notion of similarity defined in (35) is not transitive: If *a* is similar to *b* and *b* is similar to *c*, then *a* need not to be similar to *c* with the same degree as *a* to *b* and *b* to *c*. Lack of transitivity is plausible when considering similarity judgments collected, e.g., in psychological experiments. However, the interpretation of the demonstrative *so* requires a more restrictive notion of similarity. Suppose that, analogous to the example above, there are two *so*-phrases expressing that *sim*(*a*,*b*,*F*) and *sim*(*c*,*b*,*F*) where *b* is a shared demonstration target. Then *a* and *c* must also be similar, *sim*(*a*,*c*,*F*), that is, the relation has to be transitive and symmetric. This corresponds to the idea of establishing ad-hoc subkinds: If *a* and *b* belong to a subkind and *b* and *c* belong to the same subkind, then *a* and *c* must also belong to that subkind. So what we want for the similarity relation interpreting the demonstrative *so* is that it establishes an equivalence relation (reflexivity is obvious[32]).

'high-powered electric car'(*x*) = *high-powered**($\mu_{\text{horsepower}}(x)$ • c($\mu_{\text{drive\_type}}(x)$)) & *electric**($\mu_{\text{drive\_type}}(x)$)

[31] Very many thanks to one of the reviewers for providing this example.

[32] See also the argument for a non-identity implicature in footnote 14 and recall that the explanation of why definite determiners are ungrammatical in *so*-phrases (section 2.1) does not require that there is no reflexive pair

Our notion of attribute spaces offers a natural way to define similarity as an equivalence relation. Recall that an attribute space *F* comes with a set of classification functions. The set of classification functions in *F* is denoted by *C*(*F*). In (36), an equivalence relation is defined such that two individuals x and y are equivalent if and only if every classification function *p** in *C*(*F*) yields the same result when applied to the corresponding points in the attribute space, $\mu_F(x)$ and $\mu_F(y)$.

(36) $x \sim y$ iff $\forall\, p^* \in C(F): p^*(\mu_F(x)) = p^*(\mu_F(y))$

The predicate *sim* is defined as such an equivalence relation in (37):

(37) $sim(x, y, F)$ iff $\forall\, p^* \in C(F): p^*(\mu_F(x)) = p^*(\mu_F(y))$

The definition in (37) yields a strong notion of similarity: Two entities *x* and *y* are similar if and only if they agree on all elements of *C*(*F*). As already mentioned, this establishes an equivalence relation. In rough set theory (Pawlak 1998) such a relation is called *indiscernibility*. Indiscernibility of *x* and *y* basically means that, given a certain set of features, that is, a certain system of classification functions, *x* cannot be distinguished from *y* and they trigger the same inferences. Indiscernibility seems close to Nunberg's idea of *contextual granularization* (although there is no definition provided in Nunberg 2004, cf. section 2.3). It is also close to what van Rooij (2009) calls *relevant indistinguishability* in the interpretation of *same* (see also Hobbs 1985, Lasersohn 2000, and van Rooij 2011).

The question of whether similarity is an equivalence relation, and in particular the question of whether it is symmetric, is the topic of a long-standing controversy starting with Tversky's seminal paper in 1977 (cf. also Gleitmann et al. 1996). It has to be noted, however, that Tversky's findings base on the English adjective *similar,* cf. footnote 22. As mentioned in the beginning of this paper, English *similar* as well as German *ähnlich* differ in meaning from the demonstrative *so*, e.g., in being gradable. For ad-nominal *so*, however, modeling similarity as equivalence with respect to a particular set of features is adequate since it allows to view ad-nominal *so*-phrases as a way of expressing ad-hoc subkinds – being elements of the same subkind should be an equivalence. For the ad-adjectival case it may be objected that the symmetry requirement is too strong. It may be argued that the sentence *So groß ist Anna.* 'Anna is tall like this.' is true even if Anna is significantly taller than the person pointed at. The "at-least interpretation" is in fact standard for equatives – *Anna is as tall as Berta* is standardly taken to mean that Anna is at least as tall as Berta. On the other hand, the ad-adjectival *so*-phrase can be paraphrased by stating that Anna and the target of the demonstration are of the same height – *Anna has the same height as the person over there. / Anna and the person over there are of the same height*. Neither of these sentences suggests an "at-least interpretation". Although there are technical solutions available to implement an "at-least interpretation", for example by using the right closure operator, we will rather leave the question of symmetry in the case of ad-adjectival *so* as an open issue for future research.

5 Conclusion

Similarity demonstratives combine two basic modules of language and cognition, viz. demonstration and similarity. This paper can be seen as a first step into an analysis trying to account for their combination. Summarizing, we started out from the question of what similarity demonstratives like the German demonstrative *so* refer to and how it is possible that a demonstrative functions as a modifier. We hypothesized that similarity demonstratives generate ad-hoc kinds by similarity to their demonstration

but instead relies on the fact that similarity does not degenerate to identity, that is, include reflexive pairs only.

target which is an individual (or event). The focus in this paper is on ad-adjectival and ad-nominal occurrences of German *so* and on their deictic uses, excluding ad-verbal occurrences and anaphoric uses. While ad-adjectival cases express similarity with respect to only one dimension, i.e. the one provided by the adjective, ad-nominal cases express similarity with respect to a number of dimensions which have to be k-properties of the concept associated with the noun.

The semantics of the demonstrative *so* is given as a three-place similarity relation including referent, target of demonstration and a (set of) dimension(s) of similarity. Similarity is not left as a primitive relation and instead implemented with the help of multi-dimensional attribute spaces known from Artificial Intelligence. Individuals are mapped to points in such spaces by means of generalized measure functions. In the one-dimensional ad-adjectival case these functions are identical with the standard measure functions mapping individuals to metrical scales. In the multi-dimensional ad-nominal case generalized measure functions are composed out of a number of basic functions mapping individuals to scales of various types – metrical, but also ordinal and nominal ones. Attribute spaces spanned by generalized measure functions are equipped with classification functions mapping points in such spaces to truth-values. Classification functions are required to mirror predicates on individuals, that is, make the diagram in figure 1 commute, thereby linking attribute spaces back to truth-conditional semantics. Similarity as used in spelling out the semantics of German *so* is defined as indistinguishability with respect to a given set of dimensions, which is an equivalence relation.

There is a number of open questions and issues for further research. First, ad-verbal occurrences and anaphoric uses have to be addressed. This includes a closer look into the topic of dimensions associated with kinds or concepts. Another major topic addresses the properties of the "similarity as indistinguishability" relation, in particular the problem of symmetry in the ad-adjectival case. This includes comparing the analysis in this paper to analyses of sameness.

The third topic of future research addresses other expressions of similarity in German as well as across languages. Similarity demonstratives other than *so* and *solch* are German *derart* and *dermaßen* (lit: *of-this-kind* and *of-this-degree*) as well as Polish *tak* and Turkish *böyle*. Beyond demonstratives there are similarity expressions like German *ähnlich/ gleich/ dasselbe* and English *similar/ like/ the same* which differ substantially from *so* / *such* (cf. Umbach to appear). This is evidence that natural languages employ similarity relations of different types. The setting combining standard semantics with multi-dimensional attribute spaces will provide a suitable framework for the analysis and is, finally, a topic of future research on its own, regarding technical implementation as well as its role in semantics.

References

Alrenga, P. (2007) *Dimensions in the Semantics of Comparatives*. Dissertation, University of California, Santa Cruz.

Anderson, C. & Morzycki, M. (2013) Degrees as Kinds. To appear in *Natural Language and Linguistic Theory.*

Bierwisch, M. (1987) Semantik der Graduierung. In M. Bierwisch, E. Lang, E, (eds) *Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven*. Akademie Verlag, Berlin, 91-286.

Bolinger, D. (1972) *Degree Words*. Mouton, The Hague.

Carlson, G. N. (1980) Reference to Kinds in English. Garland, New York & London.

Carlson, G. (2010). Generics and Concepts. In F. J. Pelletier (ed.) *Kinds, Things and Stuff*. Oxford, Oxford University Press, 16-36.

Cresswell, M. J. (1976) The Semantics of Degree. In B. Partee (ed.), *Montague Grammar*. New York: Academic Press, 261-292.

Ehlich, K. (1986) *so* – Überlegungen zum Verhältnis sprachlicher Formen und sprachlichen Handelns, allgemein und an einem widerspenstigen Beispiel. In I. Rosengren (Hg.) *Sprache und Pragmatik*, *Lunder germanistische Forschungen* 55, 279-298.

Elbourne, P. (2009) Demonstratives as Individual Concepts. *Linguist and Philosophy* 31, 409–466

Fricke, E. (2007) *Origo, Geste und Raum – Lokaldeixis im Deutschen.* de Gruyter, Berlin / New York.

Gärdenfors, P. (2000) *Conceptual Spaces*. MIT Press.

Gleitman, L., Gleitman, H., Miller, C., Ostrin, R. (1996) *Similar*, and similar concepts. *Cognition* 58, 321-376.

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. (2003) *Manifestations of Genericity*. New York: Routledge.

Harweg, R. (1990) *Studien zur Deixis*. Universitätsverlag Dr. Norbert Brockmeyer, Bochum.

von Heusinger, K. (2011) Specificity, Referentiality and Discourse Prominence: German Indefinite Demonstratives. *Proceedings of Sinn & Bedeutung* 15, 9-30.

Hobbs, J. R. (1985) Granularity. *Proceedings of the International Joint Conference on Artificial Intelligence.*

Hole, D. & Klump, G. (2000) Definite Type and Indefinite Token: The article *son* in Colloquial German. *Linguistische Berichte* 182, 231-244.

Huddleston, R. & Pullum, G. K. (2002) The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.

Jannedy, S. (2010) The Usages and Meanings of 'so' in Spontaneous Berlin Kiezdeutsch. In M. Weirich & S. Jannedy (eds.) *ZAS Papers in Linguistics (ZASPiL) 52*, pp. 43-61.

Kaplan, D. (1989). *Demonstratives*. In J. Almog, J. Perry & H. Wittstein (eds.), *Themes from Kaplan*, Oxford University Press, 481–563.

Kennedy, C. (1999) *Projecting the adjective*. Garland Press, New York.

König, E., Stark, D., & Requardt, S. (1990) *Adverbien und Partikeln. Ein deutsch-englisches Wörterbuch*. Julius Groos Verlag, Heidelberg.

König, E. (2012) Le rôle des déictiques de manière dans le cadre d'une typologie de la deixis. *Bulletin de la Société de Linguistique de Paris* 107/1, 11-42.

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. (2009) Approximate interpretations of number words: A case for strategic communication. In E. Hinrichs & J. Nerbonne (eds.), *Theory and Evidence in Semantics*, Stanford: CSLI Publications 2009, 109-132.

Krifka, M. (2012) Definitional Generics. In A. Mari, C. Beyssade & F. Del Prete (eds.) *Genericity.* Oxford: OUP, 372-389.

Landman, M. (2006) *Variables in natural language*. Ph.D. dissertation, University of Massachusetts, Amherst.

Landman, M. & Morzycki, M. (2003*)* Event-Kinds and the Representation of Manner. In N. M. Antrim, G.Goodall, M. Schulte-Nafeh, V. Samiian (eds.)*. Proceedings of the Western Conference in Linguistics (WECOL) 2002*, vol. 11. California State University, Fresno.

Lang, E. (1989) The Semantics of Dimensional Designation of Spatial Objects. In M. Bierwisch & E. Lang (eds.), *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*. Berlin: Springer, 263- 417.

Lasersohn, P. (2000) *Same*, models, and representation. In B. Jackson, T. Matthews (eds) *Proceedings of Semantics and Linguistic Theory 10*, Ithaca: CLC Publications, 83-97.

Meier, C. (2009) A Comparative Semantics for Resemblance. In L. Kálmán (ed.) *Proceedings of the 10th Symposion on Logic and Language. LoLa10*, Hungarian Academy of Sciences, 141-148.

Moltmann, F. (2005) Comparatives without degrees: a new approach. *Proceedings of the Fifteenth Amsterdam Colloquium*, 2005, pp. 155-160.

Nunberg, G. (1993) Indexicality and Deixis. *Linguistics and Philosophy* 16: 1-43.

Nunberg, G. (2004) Indexical Descriptions and Descriptive Indexicals. In M. Reimer & A. Bezuidenhout (eds) *Descriptions and Beyond*. Oxford University Press.

Pittner, K. (1993) *So* und *wie* in Redekommentaren. *Deutsche Sprache* 21. S. 306-325.

Pawlak, Z. (1998) Granularity of knowledge, indiscernibility and rough sets. *Proceedings of the IEEE International Conference on Fuzzy Systems*, 106-110.

Prasada, S. & Dillingham, E. M. (2006) Principled and statistical connections in common sense conception. *Cognition* 99, 73-112.

Quine W. V. (1969) *Ontological Relativity and Other Essays*. New York, Columbia, University Press.

Redder, A. (1987) *wenn ... so*. Zur Korrelatfunktion von *so*. In I. Rosengren (Hg.) *Sprache und Pragmatik. Lunder germanistische Forschungen* 55, 315-326.

van Rooij, R. (2009) Adjectives, Comparison, and Measurement. Talk at the DGfS workshop on 'Comparison constructions and similarity-based classification', University of Osnabrück.

van Rooij, R. (2011) Vagueness and Linguistics. In G. Ronzitti (ed) *Vagueness: A guide.* Springer, 122-170.

Rosch, E. (1978) Principles of Categorization. In E. Rosch and B. Lloyd (eds.) *Cognition and Categorization.* Laurence Earlbaum Associates, Hillsdale, 27-48.

Sassoon, G. (2011) Adjectival vs. nominal categorization processes. *Belgian Journal of Linguistics* 25, 104-147.

Scontras, G. & Nicolae, A. (in this volume) Linkers and Modification in Tagalog.

Siegel, M. (1994) *Such*: Binding and the pro-adjective. *Linguistics and Philosophy* 17, 481-497.

Tversky, A. (1977) Features of similarity. *Psychological Review* 84: 327-352.

Umbach, C. (to appear) Expressing similarity: On some differences between adjectives and demonstratives. *Proceedings of IATL 2013, MIT Working Papers in Linguistics.*

Umbach, C. & Ebert, C. (2009) German demonstrative *so*- intensifying and hedging effects. *Sprache und Datenverabeitung* (*International Journal for Language Data Processing*) 1-2/2009, 153-168.

Wiese, H., Freywald, U., Özçelik, T., & Mayr, K. (2009) Kiezdeutsch as a Test Case for the Interaction between Grammar and Information Structure. *Interdisciplinary Studies on Information Structure* (ISIS) 12. Universitätsverlag Potsdam.

Zadeh, L.A. (1965) Fuzzy sets. *Information and Control* 8 (3), 338–353.