

# Making use of similarity in referential semantics

Helmar Gust, Carla Umbach

IKW Universität Osnabrück, ZAS Berlin  
helmar.gust@uos.de, umbach@zas.gwz-berlin.de

**Abstract.** Similarity is well-known to be a core concept of human cognition, e.g., in categorization and learning. Therefore, expressions of similarity in natural language are of special interest: How to account for their meaning including the results on similarity in Cognitive Science and Artificial Intelligence without abandoning referential semantics? In this paper we will lay out a framework connecting referential semantics to conceptual structures by generalizing the notion of measure functions known in degree semantics from the one-dimensional to the many-dimensional case mapping individuals to points in multi-dimensional attribute spaces. Similarity is then spelled out as indistinguishability with respect to a given set of attributes.

## 1 Introduction

Similarity is well-known to be basic in human cognition, e.g., in categorization and learning processes, and has given rise to a wide range of approaches in Cognitive Science and Artificial Intelligence in which the relation of similarity is captured in terms of, e.g., distance or feature bundles. In natural languages semantics linguistic expressions of similarity are of special interest: How to account for their meaning making use of the results in Cognitive Science and AI without abandoning the idea of referential semantics?

Following Umbach & Gust (2014) linguistic expressions of similarity include, in addition to adjectives like *similar* and verbs like *resemble*, demonstratives of manner, quality and/or degree like German *so*, Polish *tak* and English *such*. These demonstratives modify (some or all of) verbal and nominal and degree expressions, posing the problem of how to reconcile their demonstrative characteristics with their modifying capacity. Umbach & Gust argue that these demonstratives express similarity to the target of the demonstration gesture, and that the emerging similarity class constitutes an ad-hoc generated kind.

Similarity is spelled out with the help of multi-dimensional attribute spaces integrated into referential semantics by generalized measure functions mapping individuals to points in these spaces, generalizing the notion of measure functions familiar in degree semantics (cf. Kennedy 1999) from the one-dimensional to the many-dimensional case. Similarity is defined by the notion of indiscernability known in rough set theory (cf. Pawlak 1998) establishing an equivalence relation. Compared to Gärdenfors' conceptual spaces (Gärdenfors 2000), this approach employs a qualitative notion of similarity (as suggested by Tversky 1977) instead

of a geometric one. More importantly, while Gärdenfors' conceptual spaces are 'stand alone' systems, the approach presented here integrates a conceptual level of representation into referential semantics, and it does that in a way that has already been paved by degree semantics.

The focus of the current paper is on the formal details of this approach, which have only briefly been touched upon in Umbach & Gust (2014). The notions of attribute spaces, measure functions, representation, indiscernability and granularity will be laid out here in detail. Similarity will be defined as a categorical predicate *sim*, with a comparative relation *more\_sim* based on the categorical predicate. This notion of similarity will be compared to that in Tversky (1977) which is relational in nature. Section 2 will provide a brief summary of the linguistic data; in section 3 the basic definitions will be given; in section 4 similarity will be defined, and in section 5 this notion of similarity will be compared to Tversky's notion of similarity.

This paper grew out of a collaboration of Artificial Intelligence and Natural Language Semantics and includes both perspectives. The AI background is in knowledge representation and ontologies, feature extraction, and reasoning with examples, cf. Randall (1993). The semantic background is in referential semantics, in particular demonstratives, degree expressions and generics (Kaplan 1989, Kennedy 1999, Carlson 1980). Rather than trying to conceal the different perspectives we will make them explicit where advisable.

## 2 Demonstratives of manner, quality and degree

The approach in Umbach & Gust (2014) starts from German *so* ('such'/'like this'). It is one of a class of demonstratives found across languages that serve as modifiers of quality and/or manner and/or degree, including also, e.g., Polish *tak* and English *such*. Carlson (1980) proposed an analysis of English *such* as directly referring to a kind. This analysis was adopted by Anderson & Morzycki (2015) for Polish *tak*, which behaves analogous to German *so* in modifying nominal, verbal and also adjectival expressions, extending the notion of kinds to events and also degrees. Umbach & Gust (2014) argue that a directly kind-referring approach has a number of shortcomings and suggest an analysis based on similarity. While in the case of demonstratives like *this* the referent of the demonstrative phrase and the target of the demonstration gesture are identical – this is an in-built feature of the Kaplanian theory of demonstratives – in the case of manner/quality/degree demonstratives the referent and the target of the demonstration are similar (with respect to relevant features).

Consider the examples in (1). In (a), Anna's manner of dancing is characterized as being similar in certain respects to the dancing event the speaker is pointing at. In (b), Anna's cup is characterized as being similar to the cup the speaker is pointing at. Finally, in (c) Anna's height is characterized as being similar to the height of the person the speaker is pointing at. In all of these cases, by using the demonstrative *so* a similarity class is created based on the target of the demonstration. In the case of nouns and verbs the similarity class clearly

exhibits kind-like properties and should be considered as an ad-hoc generated kind, e.g., in (1a) there is an ad-hoc generated sub-kind of dancing events similar to the dancing pointed at, and in (1b) there is an ad-hoc generated sub-kind of cups similar to the mug pointed at. In the case of adjectives, as in (1c), it is an open issue whether the similarity class created by the use of the demonstrative should be considered as a genuine kind (see Anderson & Morzycki and Umbach & Gust; we will not go into this issue here.)

- (1) a. (speaker pointing to someone dancing):  
 So tanzt Anna auch.  
 'Anna dances like this, too.'
- b. (speaker pointing to a cup):  
 So eine Tasse hat Anna auch.  
 'Anna has such a cup / a cup like this, too.'
- c. (speaker pointing to a person):  
 So groß ist Anna auch.  
 'Anna is this tall, too.'

The most urgent question when dealing with similarity is that of the relevant respects, or features, of similarity. Without fixing relevant respects the notion of similarity would be trivial (Goodman 1972). In the case of adjectives like *tall* there is only one dimension, which is fixed by the adjective's lexical meaning — in the case of *tall* the feature of similarity is *height*. In the case of multi-dimensional adjectives like *healthy* and *beautiful*, as well as nouns and verbs, features of similarity have to be provided by the context. There are, however, constraints on which features qualify as licit in similarity comparison. Consider the anaphoric use of *so* in (2): *so ein Auto* can easily be understood as denoting a Japanese car but not as denoting a new car. These constraints are subject of an experimental study presented in Umbach (submitted) and are traced back to the idea of principled (vs. mere statistical) connections between properties and kinds discussed in the literature on generics (see Carlson 2010).

- (2) a. Anna hat ein japanisches Auto. Berta hat auch so ein Auto  
 (nämlich ein japanisches Auto).
- b. Anna hat ein neues Auto. Berta hat auch so ein Auto  
 (\*nämlich ein neues Auto).  
 'Anna has a Japanese car / a new car. Berta has such a car, too.'

One more issue when studying expressions of similarity in natural languages is the difference between demonstratives, like German *so* and English *such*, and adjectives, like German *ähnlich* and English *similar*. Although the two types of similarity expressions appear equivalent in meaning at first sight, there are fundamental differences. One of these is their behavior in additive contexts as in (3). The question-under-discussion in (3) ('Which cars do Otto and Anna drive?') has been partially answered by the preceding sentence – Otto drives a Mercedes Benz. Adding another Mercedes Benz driver should require an additive particle, which is in fact obligatory when using the demonstrative but is highly redundant when using the adjective.

- (3) (Otto drives a Mercedes Benz – what about Anna?)  
 a. Anna fährt auch so ein Auto.  
 b. ??Anna fährt auch ein ähnliches Auto.  
 ‘Anna drives such a car / a similar car, too.’

On the basis of this and other observations it is argued in Umbach (2014) that adjectives expressing similarity differ from demonstratives expressing similarity in carrying an in-built distinctiveness requirement on their arguments, which is the reason why the additive particle is not licensed in (3b). This entails that the similarity relation expressed by adjectives is irreflexive while that expressed by demonstratives is reflexive.

Similarity is, from the point of view of referential semantics, a simple predicate – *so ein Auto* / *such a car* denote an element of a set of cars similar to the target of the demonstration or antecedent (see the examples in (4,5) in section 4). Since this set exhibits kind-like properties, it is justly considered as an ad-hoc created sub-kind, for example, an ad-hoc created sub-kind of cars (cf. Umbach submitted). It would be unsatisfactory, however, if similarity were just an arbitrary predicate. The challenge posed by similarity is to gain insight into the mechanism and the constraints of this relation (and thereby into the mechanism and constraints of ad-hoc kind formation). For this reason, the similarity relation is spelled out in multi-dimensional attribute spaces inspired by knowledge representation techniques familiar in Artificial Intelligence.

In the next section, multi-dimensional attribute spaces, families of contexts, predicates on attribute spaces and convex closures thereof are defined. This is the machinery used in section 4 to implement a context-sensitive predicate *sim* and, based on that, a comparative similarity relation *more<sub>sim</sub>* relation. This is compared in section 5 to Tversky’s implementation of similarity.

### 3 Multi-dimensional attribute spaces

The basic idea of the framework presented in this paper is to have a referential semantics for natural language expressions where predicates talk about entities in the world, and a representational layer where predicates talk about abstract entities like numbers and symbols specifying attributes. Entities of the world are related to these abstract entities by processes of measurement or feature extraction or perception etc.. The representational layer facilitates comparing entities in the world with respect to their attributes and, in particular, determine whether they are similar with respect to certain attributes.

Representation by attributes is familiar in AI in, e.g., knowledge representation. Multi-dimensional attribute spaces in knowledge representation make use of attributes (i.e. dimensions) with all sorts of values, for example numbers and symbols. Points in these spaces correspond to lists of attribute-value pairs. A parallel, even if simpler conception is found in referential semantics, more precisely, in degree semantics where gradable adjectives are interpreted by measure functions mapping individuals to degrees, that is, points on a dimension with metric values. Attribute spaces can be seen as a generalization of dimensions

in degree semantics involving more than one dimension and allowing for values other than metrical ones. Likewise, mappings from individuals to points in attribute spaces can be seen as a generalization of measure functions from the one-dimensional to the many-dimensional case. From this point of view, making use of multi-dimensional attribute spaces in referential semantics is not a radical novelty but rather a generalization of the broadly accepted idea of degrees and measure functions.

The basic components of the framework presented in this paper are (i) *domains*, (ii) *attribute spaces* and *measure functions* and (iii) *predicate systems over attribute spaces*. They will be defined subsequently.

### 3.1 Domains and representations

We define a *domain* as a subset of the universe together with a set of predicates and non-overlapping sets of positive and negative examples for each predicate. We define *families of contexts* as sets of domains with related predicates that may differ, however, in their positive and negative examples. Examples must behave consistently within a family of contexts, that is, they must not change their roles.<sup>1</sup>

**Definition 1.** *Domain*

A domain is a quadruple  $\mathcal{D} = \langle D, \cdot^+, \cdot^-, P \rangle$  with:

- $D$  a set representing the domain,
- $P = \{p_1, \dots, p_n\}$  a set or family of predicates over  $D$ ,
- $\cdot^+ : \{p_1, \dots, p_n\} \rightarrow \mathcal{P}(D)$  a function which assigns positive examples to each predicate,
- $\cdot^- : \{p_1, \dots, p_n\} \rightarrow \mathcal{P}(D)$  a function which assigns negative examples to each predicate,
- $p_i^+, p_i^- \subseteq D$ .
- $p_i^+ \cap p_i^- = \emptyset$

**Definition 2.** *Family of contexts*

A family of contexts  $\mathcal{C}(P) = \{\langle D_k, \cdot^{+k}, \cdot^{-k}, P_k \rangle \mid k = 1, 2, \dots\}$  evaluates  $P$  if

- $P_k \subseteq P$
- for any two contexts  $C_i, C_j \in \mathcal{C}(P)$  and for all  $p \in P_i \cap P_j$  the following conditions hold<sup>2</sup>

---

<sup>1</sup> A side remark for semanticists: While in a classical (two-valued) truth-conditional semantics it is presumed that an individual is either  $p$  or  $\neg p$ , from the point of view of knowledge representation there may be individuals for which this is not (yet) known. This is the reason why the set of positive examples  $p^+$  need not cover the extension of  $p$  (and  $p^-$  need not cover  $\neg p$ ). The idea is that with increase of information the range of indeterminateness decreases. This behavior can be accounted for in logics with a notion of *underdefinedness* (see Muskens 1995). We will, however, not go into this issue here but rather assume that the underlying logic is classical and representations are (possibly incomplete) approximations.

<sup>2</sup> This is a slightly different formalization of NR, UD, and DD in (van Rooij 2011).

- *elements of  $p^+$  and  $p^-$  cannot change roles in a different context:*  
 $p^{+i} \times p^{-i} \cap p^{-j} \times p^{+j} = \emptyset$
- *discriminative power:  $p^{+i} \times p^{-i} \cap D_j \times D_j \neq \emptyset \rightarrow p^{+j} \times p^{-j} \neq \emptyset$*   
*(Predicates discriminating elements in one context should be able to discriminate elements in another context.)*<sup>3</sup>

Attribute spaces are common structures for representation. They generalize vector space approaches in allowing heterogeneous dimensions equipped with value sets of different scales (nominal, ordinal, interval, ratio), where value sets may themselves be attribute spaces (the values being points in such spaces).

An attribute space  $F$  is given by a set of attributes  $A = \{a_1 \dots a_n\}$ , such that for each  $a_i$  in  $A$  there is a set of values  $V_{a_i}$ . Points in an attribute space are in  $V_{a_1} \times \dots \times V_{a_n}$ . A *representation* includes an *attribute space*  $F$ , a (generalized) *measure function*  $\mu$  mapping elements of a domain into an attribute space and a set of predicates  $p^*$  talking about points in the attribute space. These predicates serve as *approximations*<sup>4</sup> of the predicates  $P$  of the domain and will be detailed in section 3.2.

**Definition 3.** *Representation*

A *representation*  $\mathcal{F} = \langle F, \mu, \cdot^*, \mathcal{D} \rangle$  of a domain  $\mathcal{D} = \langle D, \cdot^+, \cdot^-, P \rangle$  is given by

- an attribute space  $F$ ,
- a measure function  $\mu : D \rightarrow F$ ,
- $\cdot^* : P \rightarrow \Omega^F$ <sup>5</sup>

together with the consistency conditions

- $\forall x \in p^+ : p^*(\mu(x)) = \text{true}$
- $\forall x \in p^- : p^*(\mu(x)) = \text{false}$

for all  $p$  in  $P$ .

From this we get  $\mu(p_i^+) \cap \mu(p_i^-) = \emptyset$ .

### 3.2 Predicates on attribute spaces

Attribute spaces are familiar methods of representation in AI and also in some branches of natural language semantics, e.g., in frame-based approaches (Barcalou 1992). What distinguishes attribute spaces and representations as proposed in this paper is the idea of predicates on attribute spaces. On the worldly side, a domain includes a set of relevant predicates  $p \in P$ . According to the notion of a representation in this paper these predicates have counterparts on the representational side, namely the predicates  $p^* \in \cdot^*(P) = P^*$ . Counterpart predicates

<sup>3</sup> If a predicate *bike* discriminates between bikes and trikes in a context containing only bikes and trikes it should at least discriminate between bikes including trikes and cars in a context containing bikes, trikes and cars.

<sup>4</sup> More precisely:  $p^* \circ \mu$  approximates  $p$

<sup>5</sup> where  $\Omega^F$  is the set of characteristic functions in  $F$

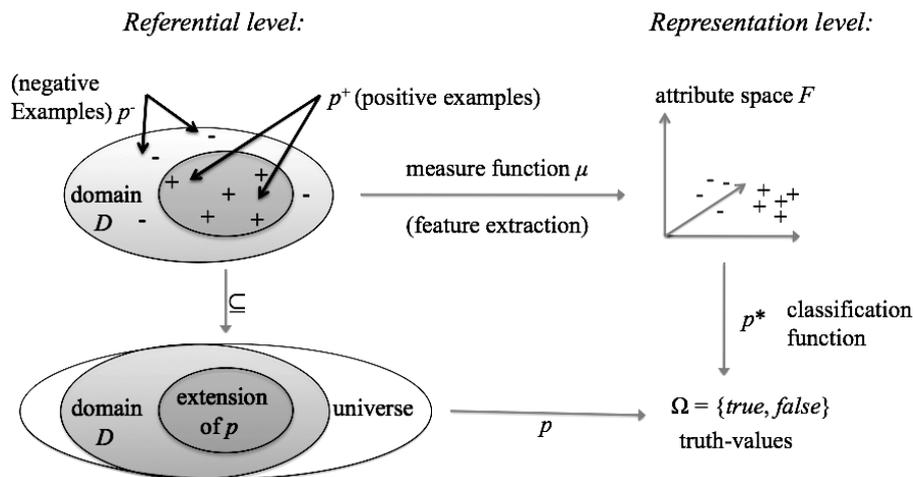


Fig. 1. Domains and Representaions

are required to be consistent with their originals; more precisely, they have to agree in truth value with the set of positive / negative exemplars of the original predicate (see definition 3). Moreover, counterpart predicates will be assumed to be convex, that is, true of all points in the convex closure of (the images of) the positive exemplars (see below).

Reasons for furnishing attribute spaces with predicates are, (i) they facilitate a straightforward definition of *indiscernability* in the sense of rough set theory (Pawlak 1998), that is, a relation establishing classes of points which count as equivalent (in a specific context, with respect to specific inferences). The indiscernability relation provides attribute spaces with a grid of granularity, allowing comparison of attribute spaces of different granularity which are identical otherwise. Indiscernability and granularity will be exploited to define a categorical and a comparative notion of similarity in the next section.<sup>6</sup>

From the point of view of AI and knowledge representation there is another reason to have predicates on attribute spaces. If you assume that elements in a domain  $D$  are not directly accessible and instead require some sort of feature extraction or perception or measuring processes, then reasoning is possible only with the results of such processes, that is, on the level of representations.

**Convex closures on attribute spaces** It has been argued by Gärdenfors (2000) that natural properties correspond to convex regions in conceptual spaces,

<sup>6</sup> It has been objected that in order to account for different granularity of representations one could use different sets of attributes and values, that is, different attribute spaces. This would require, however, a way to compare granularity directly on attribute spaces, which would be less intuitive and in the end close to what we suggest.

convexity being required by cognitive economy in learning and memory. Since Gärdenfors' conceptual spaces are metrical there is a notion of *betweenness* facilitating a definition of convexity such that a region is convex if for all points  $x$  and  $y$  in that region any point  $z$  between  $x$  and  $y$  is also in that region. Attribute spaces as specified in this paper do not have a metrical *between* relation. Convexity in these spaces will be defined by convex closure operators.

**Definition 4.** *Closure operator on attribute spaces*

An operator  $cl$  is a closure operator on an attribute space  $F$  iff

- $cl : \mathcal{P}(F) \rightarrow \mathcal{P}(F)$
- For all  $X \subseteq F$ 
  - $X \subseteq cl(X)$
  - $cl(cl(X)) = cl(X)$
- $cl(\emptyset) = \emptyset$

**Definition 5.** *Convex closure*

A closure is convex if the anti-exchange property holds:

$$y, z \notin cl(X) \wedge y \neq z \wedge z \in cl(X \cup \{y\}) \rightarrow y \notin cl(X \cup \{z\})$$

We assume, as does Gärdenfors, that predicates in a domain are well-behaved in the sense that they cover convex regions, that is in our case, they correspond to convex subsets of an attribute space. More precisely, predicates  $p^*$  on attribute spaces are required to be true for every point in the convex closure of the predicate's positive examples when mapped to the attribute space. This is, we expect representations to be *strongly consistent*.

**Definition 6.** *Strong consistency*

A representation  $\mathcal{F} = \langle \langle F, cl \rangle, \mu, *, \mathcal{D} \rangle$  of a domain  $\mathcal{D} = \langle D, .^+, .^-, P \rangle$  is strongly consistent if  $\forall p_i \in P : p_i^*(cl(\mu(p_i^+))) = \{true\}$ <sup>7</sup>

We can make this even stronger by stipulating that the  $p_i^*$  itself have convex extensions, which implies strong consistency.

Above we defined a number of constraints on closure operators. That does not, however, answer the question of how to define a particular closure operator. One way of doing that is by exploiting a partial order. On a (partially) ordered set  $\langle D, \leq \rangle$  we can define closure operators in a natural way:

**Definition 7.** *Closure operations on partial orders*

For  $A \subseteq D$  we define:

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

<sup>7</sup> We make use of the convention that a function  $f$  applied to set of arguments refers to the set of images:  $f(X) = \{f(x) \mid x \in X\}$ .

For all three constructions the anti-exchange property holds, so all the resulting sets are convex. Thus, if  $F$  is (partially) ordered, we can easily define a convex closure operator. In order to emphasize the aspect that an attribute space  $F$  is equipped with a closure operator  $cl$ , we add it to the representation:  $\mathcal{F} = \langle \langle F, cl \rangle, \mu, \cdot^*, \mathcal{D} \rangle$  of a domain  $\mathcal{D} = \langle D, \cdot^+, \cdot^-, P \rangle$ .

**Predicate systems** Predicates defined on attribute spaces classify points in such spaces facilitating different levels of granularity of representation. Given a set of base predicates  $B$  (e.g. *small*, *medium*, *big*) over a feature space  $F$  we define a set of predicates  $\tilde{B}$  inductively (analogous to defining a topology relative to a base):

- $B \subseteq \tilde{B}$
- $X \in \tilde{B} \wedge Y \in \tilde{B} \rightarrow X \cap Y \in \tilde{B}$
- $X \in \tilde{B} \wedge Y \in \tilde{B} \rightarrow cl(X \cup Y) \in \tilde{B}$

If  $F$  is (partially) ordered we get additionally

- $X \in \tilde{B} \rightarrow cl_{\rightarrow}(X) \in \tilde{B}$
- $X \in \tilde{B} \rightarrow cl_{\leftarrow}(X) \in \tilde{B}$

We assume that the elements of  $B$  are convex ( $cl(X) = X$  for  $X \in B$ ) and we say that  $\tilde{B}$  is generated by  $B$  and  $cl$ . We can view  $\tilde{\cdot}$  as operator generating the predicate system for a give base. A good candidate for  $B$  is  $\{p_1^*, \dots, p_n^*\}$ . For a representation  $\mathcal{F} = \langle \langle F, cl \rangle, \mu, \cdot^*, \langle D, \cdot^+, \cdot^-, P \rangle \rangle$ ,  $\tilde{P}^*$  is the predicate system generated by  $P^* = \{p_1^*, \dots, p_n^*\}$  and  $cl$ .

**Indiscernability** Two elements in a domain  $D$  are indistinguishable on the representational level if they are mapped by  $\mu$  onto the same point of the attribute space. But even if they are mapped to different points, they may nevertheless be indistinguishable in yielding the same inferences. This is the case if they agree on all predicates. We borrow the term *indiscernable* from Rough Set Theory:

**Definition 8.** *Indiscernable*

Given a representation  $\mathcal{F} = \langle F, \mu, \cdot^*, \langle D, \cdot^+, \cdot^-, P \rangle \rangle$  we define:

$$x \sim_{\mathcal{F}} y \equiv \forall q \in \tilde{P}^* : q(x) \leftrightarrow q(y)$$

**Granularity** For two representations  $\mathcal{F}$  and  $\mathcal{F}'$  sharing the same attribute space  $F$  we can ask whether one is more fine grained than the other, that is, whether two entities can be distinguished in one representation but not in the other one. Since indiscernability of entities in a representation depends on the set of predicates  $P$  provided by the domain, granularity of representations depends on the set of predicates  $P$ , too.

**Definition 9.** *Family of representations evaluating a system of predicates*

Given a family of contexts  $\mathcal{C}(P) = \{ \langle D_k, \cdot^{+k}, \cdot^{-k}, P_k \rangle \mid k = 1, 2, \dots \}$  evaluating a system of predicates  $P$  we get the corresponding family of representations

$$\mathcal{FC}(P) = \{ \langle F, \mu, \cdot^{*k}, \langle D_k, \cdot^{+k}, \cdot^{-k}, P_k \rangle \rangle \mid k = 1, 2, \dots \}$$

On  $\mathcal{FC}(P)$  we can define a partial order:

**Definition 10.** *Coarser representation*

Given two representations

$$\mathcal{F}_i = \langle F, \mu, \cdot^{*i}, \langle D_i, \cdot^{+i}, \cdot^{-i}, P_i \rangle \rangle \in \mathcal{FC}(P) \text{ and}$$

$$\mathcal{F}_j = \langle F, \mu, \cdot^{*j}, \langle D_j, \cdot^{+j}, \cdot^{-j}, P_j \rangle \rangle \in \mathcal{FC}(P)$$

we define:

$$\mathcal{F}_j \geq \mathcal{F}_i \text{ iff } P_j \subseteq P_i \wedge \forall x, y \in F : x \sim_{\mathcal{F}_i} y \rightarrow x \sim_{\mathcal{F}_j} y$$

We say  $\mathcal{F}_j$  is coarser than  $\mathcal{F}_i$  ( $\mathcal{F}_j > \mathcal{F}_i$ ) iff  $\mathcal{F}_j \geq \mathcal{F}_i$  but not  $\mathcal{F}_j \leq \mathcal{F}_i$ .

This completes the formal machinery required for the definition of similarity in the next section.

## 4 Defining Similarity

### 4.1 Similar

We started out from the analysis of demonstratives of manner, quality and degree, in particular German *so*, in Umbach & Gust (2014). It was argued that these demonstratives express similarity to the target of the demonstration gesture, and that the emerging similarity class constitutes an ad-hoc generated kind. This suggests that the similarity relation expressed by the demonstratives is adequately captured by the notion of indiscernability as defined above.

**Definition 11.** *Similar*

$$\forall x, y, \in D : \text{sim}(x, y, \mathcal{F}) \equiv \mu(x) \sim_{\mathcal{F}} \mu(y)$$

Definition 11 entails that the notion of similarity expressed by the demonstratives is an equivalence relation (see also section 5).

The *sim* predicate is used in Umbach & Gust (2014) for a compositional interpretation of the demonstrative *so*. When modifying nominals *so* occurs in an ad-determiner position and is thus combined with the indefinite determiner, see (4).

In the nominal case (and in the verbal case) features of comparison have to be inferred from the context of the utterance – *so eine Tasse* ('such a cup') may denote cups similar in form and/or color and/or design etc. – whereas in the adjectival case there is only one feature of comparison, which is fixed by the lexical meaning of the adjective – *so groß* ('this tall') can only mean similar in height.

We assume that features of comparison provide attributes spanning the attribute space in a representation  $\mathcal{F}$ . We write  $\mathcal{F}(f)$  to indicate a representation with a single linguistically fixed attribute  $f$ . The variable  $t$  in (4,5) is a free variable denoting the target of the demonstration gesture.

$$\begin{aligned} (4) \text{ so eine Tasse} & \quad (\text{'such a cup'}) \\ [[\text{so}]] & \quad = \lambda D. \lambda P. D(\lambda x. \text{sim}(x, t, \mathcal{F}) \wedge P(x)) \\ [[\text{so ein}]] & \quad = \lambda P. \lambda Q. \exists x. \text{sim}(x, t, \mathcal{F}) \wedge P(x) \wedge Q(x) \\ [[\text{so eine Tasse}]] & \quad = \lambda Q. \exists x. \text{sim}(x, t, \mathcal{F}) \wedge \text{cup}(x) \wedge Q(x) \end{aligned}$$

$$\begin{aligned}
 (5) \text{ so groß} & \quad (\text{'this tall'}) \\
 [[\text{so}]] & = \lambda f. \lambda x. \text{sim}(x, t, \mathcal{F}(f)) \\
 [[\text{so groß}]] & = \lambda x. \text{sim}(x, t, \mathcal{F}(\text{height}))
 \end{aligned}$$

The predicate *sim*, which is a simple modifier of determiners and degree expressions in (4,5), receives its explanatory power by being defined as indiscernability in attribute spaces.

## 4.2 More Similar

Similarity as expressed by demonstratives is clearly not gradable – there is no way of saying that something is *more such a thing* than something else. Thus the *sim* predicate defined above is appropriate for the semantics of, e.g., English *such* and German *so*. In contrast, similarity as expressed by adjectives is gradable, cf. English *more similar* and German *ähnlicher*. We will provide a gradable notion of similarity in the way of the vague predicate analysis of Klein (1980) such that the comparative is based on the positive. We define a four-place predicate  $\text{more\_sim}(x, y, z, u, \mathcal{F}\mathcal{C}(P))$  with the help of representations of different granularity in a family of representations such that  $x$  is more similar to  $z$  than  $y$  to  $u$ , if there is a representation  $\mathcal{F}'$  identifying  $x$  and  $z$  while distinguishing  $y$  and  $u$ . Any  $\mathcal{F}''$  identifying  $y$  and  $u$  must not discriminate  $x$  and  $z$ .

### Definition 12. More similar

Given a family of representations  $\mathcal{F}\mathcal{C}(P)$ ,  $x$  is more similar to  $z$  than  $y$  to  $u$ ,  $\text{more\_sim}(x, y, z, u, \mathcal{F}\mathcal{C}(P))$  iff

$$\begin{aligned}
 & \exists \mathcal{F}' \in \mathcal{F}\mathcal{C}(P) : \text{sim}(x, z, \mathcal{F}') \wedge \neg \text{sim}(y, u, \mathcal{F}') \\
 & \wedge \forall \mathcal{F}'' \in \mathcal{F}\mathcal{C}(P) : \text{sim}(y, u, \mathcal{F}'') \rightarrow \text{sim}(x, z, \mathcal{F}'')
 \end{aligned}$$

The 4-place predicate  $\text{more\_sim}$  provides an interpretation for natural language comparatives as in *Anna resembles her mother more than Berta resembles her father*. Definition of a 3-place version is straightforward:

$$\text{more\_sim}_3(x, y, z, \mathcal{F}\mathcal{C}(P)) \text{ iff } \text{more\_sim}(x, y, z, z, \mathcal{F}\mathcal{C}(P)).$$

Another motivation for a comparative version of similarity stems from multi-dimensional gradable adjectives like *healthy* and *beautiful* which are difficult to handle in standard degree semantics. One-dimensional adjectives like *tall* are interpreted in degree semantics as measure functions mapping individuals to degrees (Kennedy 1999). The positive form (as in *A is tall*) is assumed to include a context-dependent cut-off degree for individuals to count as tall in the context/comparison class. For many-dimensional adjectives, however, it is unclear how to define a cut-off. Sassoon (2013) suggests quantification over dimensions such that, e.g., *healthy* means *healthy in all relevant dimensions* while leaving the question open of how to define the comparative.

Umbach (in press) makes use of similarity in multi-dimensional spaces when interpreting evaluative predicates like *beautiful*. Since evaluative predicates don't have a linguistically fixed denotation there is no independently given set of positive exemplars  $\text{beautiful}^+$ . Whether something is beautiful or not is not a matter of fact but rather a matter of convention, that is, of negotiation of criteria. Negotiating criteria can be seen as determining a 'prototype' which is, however,

not a point but rather a predicate on points in an attribute space. Suppose *proto\_beautiful* is the class of points in  $F$  corresponding to the criteria for something to count as beautiful, that is, *proto\_beautiful* plays the role of a cut-off. The counterpart predicate *beautiful\** then has to be the convex closure of *proto\_beautiful*. An individual can now be said to count as beautiful in a domain  $\mathcal{D}$  and a representation  $\mathcal{F}$  iff it is indiscernable from the elements of *beautiful\**, that is, the closure of *proto\_beautiful* (where the domain serves as comparison class).

$\forall x \in \mathcal{D}$ :  $x$  is beautiful in  $\mathcal{F}$  iff  
*beautiful\**( $\mu(x)$ ) in  $\mathcal{F}$ , that is,  $cl(proto\_beautiful)(\mu(x))$  in  $\mathcal{F}$ .

The comparative of multi-dimensional adjectives is interpreted with the help of the *more\_sim* relation.<sup>8</sup>

$\forall x, y \in \mathcal{D}$ :  $x$  is more beautiful than  $y$  in  $\mathcal{FC}(P)$  iff  
*more\_sim*<sub>3</sub>( $x, y, cl(proto\_beautiful), \mathcal{FC}(P)$ )<sup>9</sup>

## 5 Comparison to Tversky's account of similarity

Tversky's contrast model of similarity was developed as an alternative to the at that time dominant geometric models in which (dis)similarity of two objects is represented by distance. Tversky started from empirical findings that appeared incompatible with the axioms of a metric distance function. He claimed that " ... the assessment of similarity between stimuli may better be described as a comparison of features rather than as the computation of metric distance between points." (p. 328), and proposed to assess similarity making use of sets of features (binary, nominal, ordinal, cardinal). Similarity of two objects  $a$  and  $b$  is measured by a matching function  $F$  accounting for communalities of and differences between their feature representations  $A$  and  $B$ :

$$sim(a, b) = F(A \cap B, A - B, B - A).$$

The result of the matching function corresponds to a value on an interval scale  $S$  which makes it possible to compare the similarity of two objects to that of two other objects:  $sim(a, b) \leq sim(c, d)$  iff  $S(a, b) \leq S(c, d)$ .<sup>10</sup>

From the point of view of the semantics of gradable adjectives, Tversky's account of similarity is in the spirit of degree semantics where the comparative is prior to the positive. Tversky's notion of similarity maps pairs of individuals to degrees

<sup>8</sup> This corresponds to a vague predicate interpretation of multi-dimensional adjectives. It's an open question, however, whether one-dimensional adjectives like *tall* should be interpreted in this way, too. One of our future issues is to exploit a two-way approach: While the comparative of adjectives relating to a single metrical scale makes use of that scale (in a degree semantics fashion), the comparative of multi-dimensional adjectives involves similarity to prototypes, as sketched above.

<sup>9</sup> We use *more\_sim*<sub>3</sub> also with arguments from  $F$ . This is unproblematic because  $\mu$  mediates between the two levels.

<sup>10</sup> where  $S$  computes the values of a model based on the weighted sums over common and distinctive features, which means, there exist  $f, \theta, \alpha, \beta$  with  $S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$ .

on a scale of similarity thereby facilitating comparison — *a and b are more similar to each other than c and d*. Positive judgments, as in *a and b are similar*, are not considered. In contrast to Tversky’s account, the notion of similarity developed in this paper is in the spirit of the vague predicate analyses starting from the positive form and defining the comparative ‘on top’.

Tversky started from empirical findings which seem incompatible with the axioms of a metric distance function, i.e. minimality, symmetry and triangle inequality. Minimality appears problematic if identification probability is interpreted as a measure of similarity because identical stimuli are not always identified by subjects as being identical. Triangle inequality is hardly compelling, following Tversky, in view of cases involving different features. For example, Jamaica is similar to Cuba (with respect to geographical proximity) and Cuba is similar to Russia (with respect to political affinity) but Jamaica and Russia are not similar at all. Finally, symmetry is apparently false considering experimental results showing that the judged similarity of North Korea to Red China exceeds the judged similarity of Red China to North Korea.

In this paper, similarity is viewed as a means of classification, that is, a relation establishing classes, or kinds, of individuals. From a classification perspective, similarity has to be reflexive, symmetric and transitive. Concerning reflexivity, Umbach (2014) shows that the interpretation of demonstratives like German *so* and English *such* requires reflexivity while the interpretation of adjectives expressing similarity, like German *ähnlich* and English *similar*, excludes reflexives pairs. Being irreflexive explains why adjectival similarity expressions lack a kind-forming capacity — kinds have to be equivalence classes.

Concerning symmetry, Gleitman et al. (1996) present a series of studies showing that Tversky’s results on similarity judgments are due to figureground effects in presentation and argue that similarity is a genuinely symmetric relation. In fact, Tversky himself already suggested that non-directional similarity statements (*North Korea and Red China are similar*) are symmetric. Finally, transitivity is obvious in the case of demonstratives. Consider the sentence *Anna hat so ein Auto und Berta hat so ein Auto*. (‘Anna has a car like this and Berta has a car like this.’) where the two *so*-phrases share the same demonstration target. This sentence clearly entails that Anna’s and Berta’s cars belong to the same subkind of cars. For adjectives expressing similarity the situation is again less clear. It might be possible that changing the respects of similarity, as in the Jamaica/Cuba/Russia example, is licensed by adjectives (the sentence *Jamaica is similar to Cuba and Cuba is similar to Russia but Jamaica and Russia are not similar* appears acceptable in spite of different features of similarity).

Summing up, similarity as expressed by natural language demonstratives (e.g., German *so* and English *such*) is in fact an equivalence relation whereas similarity as expressed by adjectives (e.g., German *ähnlich* and English *similar*) is not. This confirms the analysis that similarity demonstratives, but not similarity adjectives, ad-hoc establish novel (sub-)kinds.

## 6 Summary

In this paper, a representational framework was presented featuring multi-dimensional attribute spaces equipped with systems of convex predicates. This framework facilitates the definition of a notion of similarity, or indiscernability, as required in classification processes, which is suited for the interpretation of demonstratives of manner, quality and degree in natural language. In (Umbach in press), a similar approach was used for evaluative propositions.

The framework emerged out of a collaboration of Artificial Intelligence and Natural Language Semantics. From an AI point of view it can be seen as a means of representing (possibly incomplete) knowledge about a given domain. From a semantic point of view it can be seen as extending referential semantics in order to integrate conceptual aspects of meaning.

## References

1. Anderson, C., Morzycki, M. (2015): Degrees as Kinds. To appear in *Natural Language and Linguistic Theory* 2015.
2. Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (eds.), *Frames, fields, and contrasts*. Hillsdale, NJ: Lawrence Erlbaum Associates.
3. Carlson, G. N. (1980): *Reference to Kinds in English*. Garland, New York & London.
4. Carlson, G. (2010): Generics and Concepts. In F. J. Pelletier (ed.) *Kinds, Things and Stuff*. Oxford, Oxford University Press, 16-36.
5. Gärdenfors, P. (2000): *Conceptual Spaces*. MIT Press.
6. Gleitman, L., Gleitman, H., Miller, C., Ostrin, R. (1996): Similar, and similar concepts. *Cognition* 58, 321-376.
7. Goodman, N. (1972). Seven strictures on similarity. In Goodman, N. (ed.) *Problems and Projects*. The Bobbs Merrill Company, Indianapolis and New York, 437-447.
8. Kaplan, D. (1989): Demonstratives. In J. Almog, J. Perry & H. Wittstein (eds.), *Themes from Kaplan*, Oxford University Press, 481-563.
9. Kennedy, C. (1999): *Projecting the adjective*. Garland Press, New York.
10. Klein, E. (1980): A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4:1-45.
11. Muskens, R. (1995): *Meaning and Partiality*. CSLI, Stanford.
12. Pawlak, Z. (1998): Granularity of knowledge, indiscernibility and rough sets. *Proceedings of the IEEE International Conference on Fuzzy Systems*, 106-110.
13. Davis, Randall; Howard Shrobe; Peter Szolovits (Spring 1993). "What Is a Knowledge Representation?". *AI Magazine* 14 (1): 17-33.
14. van Rooij, Robert (2011): Measurement and Interadjective Comparisons. *Journal of Semantics*, 28, 2011: 335-358.
15. Sassoon, G. (2013): A Typology of Multidimensional Adjectives. *Journal of Semantics* 30: 335-380.
16. Tversky, Amos (1977): Features of similarity. *Psychological Review* 84, 327-352.
17. Umbach, C., Gust, H. (2014): Similarity demonstratives. *Lingua* 149: 74-93.
18. Umbach, C. (2014): Expressing similarity: On some differences between adjectives and demonstratives. *Proc. of IATL 2013, MIT Working Papers in Linguistics*.
19. Umbach, C. (in press): Evaluative propositions and subjective judgments. In J. van Wijnbergen-Huitink & C. Meier (eds.) *Subjective meaning*. Berlin, de Gruyter.
20. Umbach, C. (submitted): Ad-hoc kind formation by similarity.