

Expressing Similarity

– The semantics of linguistic expressions of similarity

Research project Oct 2012 - Sept 2015, financed by the Deutsche Forschungsgemeinschaft DFG

Research area: General Linguistics, Semantics, Pragmatics

Principal investigator: Dr. habil. Carla Umbach

Affiliation: Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin

1. Summary

This project investigates linguistic expressions of similarity. The paradigm lexical item in German expressing similarity is the demonstrative *so*. It can be used deictically as well as anaphorically, and in addition plays a central role in German equative constructions – scalar as well as non-scalar ones. Until now, no satisfactory semantic analysis has been suggested for the demonstrative *so* accounting for its deictic and anaphoric uses and, at the same time, for its role in equative comparison.

In Cognitive Science similarity is a fundamental concept in explaining human perception and classification. It will be shown in the project that similarity is fundamental in the semantics of natural language, too. Combining standard truth-conditional semantics with techniques from Artificial Intelligence, an interpretation of the demonstrative *so* will be given, based on similarity as a basic semantic relation.

Starting from the demonstrative *so* the analysis will be extended to other expressions of similarity (e.g. *ähnlich* 'similar', *genauso / gleich* 'same'), making use of a range of similarity measures. Cross-linguistic differences between German, English, Spanish and Dutch will be investigated looking for language-specific strategies in expressing similarity.

2. State of the art and preliminary work

Languages provide multiple ways to express similarity, which is no surprise taking into account that similarity is a basic feature in human cognition, essential for explaining the classification of novel objects and situations. The key objective of this project proposal is a semantic analysis of expressions of similarity, combining findings and methods from standard truth-conditional semantics with results and techniques from Cognitive Science and Artificial Intelligence. It will be shown that similarity is a fundamental semantic relation, next to identity, manifest in a variety of linguistic expressions.

The paradigm lexical item in German expressing similarity is the **demonstrative** *so*. Although it occurs in a broad range of uses,¹ its core meaning is **deictic** or **anaphoric**, relating to "aspects of objects" (Ehlich 1987). While there exists extensive literature on the non-deictic/-anaphoric uses, the semantic interpretation of the deictic/anaphoric cases is an as yet unresolved issue, the core problem being the question of how to explain reference to aspects of an object instead of the object itself.

The deictic and the anaphoric use of the demonstrative will be the starting point of the proposed project. Deictic uses are accompanied by a demonstration gesture, anaphoric uses require an

¹For an overview see König et al. 1990. Recent research is focused on particular uses, for example, the contraction of *so* and the indefinite article *ein* (*Kaufst du mir so'n Pullover?* 'Will you buy me a jumper like this?', cf. Hole & Klump 2000), and its relation to indefinite *this* in English (*Eva will so'n Film über Eliade sehen. / Eva wants to watch this movie about Eliade.*, von Heusinger 2010). 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'). König (2010) investigates manner deixis from a typological point of view, including German *so* (*So macht man das.* 'That's the way to do it.').

antecedent.² The examples in (1) show deictic uses, combined 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 either be gradable or non-gradable.³ Combinations with gradable expressions are called **scalar**, those with non-gradable ones are called **non-scalar**. The examples in (2) show anaphoric uses. In the project, an analysis will be developed combining results from the research on demonstratives (cf. section a), and the research on the meaning of the English anaphor *such* (section b), making use of the notion of similarity in Cognitive Science and Artificial Intelligence (section c).

(1) deictic

- | | |
|--|--------------------------------------|
| a. (speaker pointing to a person): So groß ist Anna. | 'Anna is that tall.' |
| b. (speaker pointing to a car): So ein Auto hat Anna. | 'Anna has a car like this.' |
| c. (speaker pointing to someone dividing a fish):
So hat Anna den Fisch auch zerlegt. | 'Anna cut the fish like that , too.' |

(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 paradigm grammatical construction expressing similarity is **equative comparison**, cf (3), which is traditionally interpreted 'en bloc', with *so* and *wie* as mere connecting pieces. The current proposal is based on the hypothesis that the expression *so* in German equatives is an instance of the demonstrative *so* and has to be interpreted analogous to the deictic and anaphoric uses. This will allow for a fully compositional analysis such that the fact that equatives express similarity results from the meaning of the demonstrative.⁴ Moreover, it will facilitate the analysis of non-scalar equatives, i.e. those based on non-gradable expressions as in (3b), for which no semantic interpretation is available up to now. The analysis of equatives will make use of the wealth of results on gradability and comparison (section d) and of detailed descriptive studies on comparison constructions in German (section e).

(3) equatives

- | | |
|---|--------------------------------------|
| a. Anna ist so groß wie Berta. | 'Anna is as tall as Berta.' |
| b. Anna hat so ein Auto wie Berta. | 'Anna has the same car as Berta.' |
| c. Anna hat den Fisch so zerlegt wie Berta. | ' Anna cut the fish like Berta did.' |

The analysis of the demonstrative *so* will be extended including **other German expressions of similarity**, for example *solch* 'such', *genauso* 'exactly as', *gleich* 'same', and *ähnlich* 'similar', which can be used anaphorically and in equative-like constructions, cf. (4a,b). The question of how the interpretation of such expressions compares to that of the demonstrative *so* will lead to a more refined picture of similarity expressions, making use of a range of different similarity measures suggested in the Cognitive Science literature (cf. section c). There is a number of papers investigating the semantics of English *same* and *similar* to be exploited in the project (section f).

(4) other similarity expressions

- | |
|--|
| a. (Berta ist 180.) Anna ist genauso groß /ähnlich groß (wie Berta).
'(Berta is 180.) Anna is exactly as tall / similarly tall (as Berta).' |
| b. (Berta hat ein Auto mit Ladeklappe.) Anna hat genauso ein / auch solch ein / ein ähnliches Auto (wie Berta).
'(Berta has a car with a hatch.) Anna has also such a / the same / a similar car (like Berta).' |

²The intensifier use of *so*, in (a), and the hedging use, in (b), differ from deictic/anaphoric uses in occurring "out of the blue", without gesture or antecedent, cf. Umbach & Ebert (2009).

(a) <i>Die Pizzeria Aroma ist so teuer.</i>	'The pizzeria Aroma is so expensive.'
(b) <i>Der Bischof hatte so'ne Mütze auf.</i>	'The bishop had something like a cap on his head.'

³ Examples for gradable nouns and verbs are *Geizhals* ('nickel nurser') and *rasen* ('speed').

⁴ Although close in meaning to equatives, result clauses –*Bertha ist so groß, dass sie an das Regal kommt*. 'Berta is sufficiently tall to reach the board.', cf. Meier (2001) – will not be considered in the current proposal.

The examples in (1)-(3) suggest that, in German, there is a close correspondence between the deictic/anaphoric use of the demonstrative *so* and its occurrence in equative constructions, and there is also a close correspondence between scalar and non-scalar comparison. English exhibits neither of these correspondences, whereas Spanish shows part of them and Dutch appears close to German (cf. the table in section 3.2, P6). The proposed project will, in addition to the investigation of similarity expressions in German, include a **cross-linguistic component** contrasting the results from German with data from English, Spanish and Dutch in order to find out, how the strategies of expressing similarity in these languages differ from the ones found for German (cf. section g).

The theoretical background of this project is in truth-conditional semantics complemented by concepts from Cognitive Science and Artificial Intelligence. Linguistic data will be obtained by in-depth corpus studies (using mono-lingual as well as multi-lingual parallel corpora), supplemented by questionnaires and experimental elicitation of data.

2.1 Description

(a) The meaning of demonstratives

Following Ehlich (1987), the expression *so* is, first of all, a demonstrative picking up "aspects of objects" which are used to modify the denotation of adjectives, nouns and verbs. In (1a) the height of the person the speaker points to is used to characterize Anna's height. In (1b) certain properties of the car the speaker points to are used to characterize Anna's car. In (1c) the manner of the fish-cutting event the speaker points to is used to characterize Anna's way of doing that. This raises two questions, (i) what does the demonstrative *so* refer to?, and (ii) how is it possible that a demonstrative acts as a modifier? One readily available answer consists in assuming that the demonstrative refers to properties. That would mean, however, to employ a semantic framework based on property theory (cf., e.g., Chierchia & Turner 1988). In the present proposal this option will be dismissed because it shifts the explanatory burden to the semantic framework. We will presuppose a standard semantic ontology, including individuals and events, but no properties as such, which leaves us with the problem of the referent of the demonstrative *so* – if you cannot refer to degrees or properties or manner, **what does the speaker refer to** in the examples in (1a-c)?

This problem must not be confused with Quine's negative view on reference in general, arguing that reference is inherently indeterminate (Quine 1960). Our problem in, e.g., (1b) is not to determine the object the speaker points to – we know that it is the car because *so* is combined with *Auto* ('car') in (1b). It is unclear, however, how the demonstrative *so* picks up properties of the object, and moreover, it is unclear which properties it picks up.

The standard theory of demonstratives is the direct reference theory, according to which certain singular terms refer directly, without the mediation of a Fregean Sinn (cf. Kaplan 1989). The most important consequence of this idea is the separation of the character and the content of an expression, the character depending on the context of utterance, and the content depending on world and time of evaluation.

Nunberg (1993, 2002) proposes an elaboration of Kaplan's theory addressing the problem of so-called **descriptive uses**, in which the object referred to is not identical to the interpretation of the indexical or demonstrative.⁵ Nunberg suggests dismissing the assumption that the referent is identical to the interpretation of the indexical or demonstrative, and allowing for other relations in addition. Following Nunberg, the semantics of a demonstrative involves (i) a deictic component picking out a referent, (ii) an interpretation contributed to the proposition, and (iii) a relation between the referent and the interpretation. This theory offers a solution to the problem of the referent of the demonstrative *so*, which will be one of the central hypotheses of the present proposal, in short: The referent of the demonstrative *so* is an individual or event. Its meaning, however, consists in establishing a **similarity relation** between the referent pointed to and the referent of the phrase modified by the demonstrative – for example, in (1b) between the car the speaker points to and Anna's car. For this idea to be

⁵ For example, if someone utters *He is usually an Italian*, while pointing at Benedict XVI, the sentence will not mean that Benedict XVI is usually an Italian, and instead, that the person holding the position of the pope is usually an Italian.

productive, we need a notion of similarity which is not a primitive, and instead endowed with characteristics gained from research in similarity in Cognitive Science.

(b) English *so* and *such*

English *so* and *such* are close in meaning but not equivalent, to German *so* and *solch*. First, English *such*, like German *solch*, can only be combined with nominal expressions. Secondly, English *so*, unlike German *so*, is restricted to combination with adjectives and verbs. Thirdly, English *so* and *such* have deictic and anaphoric uses, like their German counterparts, but English *so* requires negation or question contexts when occurring in equative comparison (*The old lady is not so innocent as she seems*). Finally, different from German, English *so* is frequently found as a clausal pro-form (*She thought he was wrong but was too polite to say so*), cf. Cornish (1992) and Kehler & Ward (1999). For a comprehensive overview see Huddleston & Pullum (2002), chap. 13 and 17.

The most prominent study on English *so* and *such* is Bolinger's (1972) book on degree words. Bolinger distinguishes the **identifier** use of English *so* and *such* from their **intensifier** use. The identifier use is deictic or anaphoric (*He is about so tall. – We need a telescope equipped for solar photography. Such a telescope is hard to find*). The intensifier use occurs in result clauses or out of the blue, like the intensifier use of German *so* (cf. footnote 2) (*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 contains 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*).⁶ Carlson considers these problems as evidence that *such* relates to a **kind** instead of a modifier and suggests a semantic interpretation introducing a free kind variable denoting a subtype of the modified common noun. Siegel (1994) argues, contra Carlson, that *such* is a simple **pro-adjective** which is not bound by an adjective (or other nominal modifier), and 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 extends Carlson's analysis to the paraphrase *like that* and to adnominal German *so* and Polish *tak*. Landman moreover suggests an interpretation of the adverbial use of German *so* and Polish *tak* by postulating event-kinds as an ontological category (see also Landman & Morzycki 2003).

The approaches by Carlson, Siegel and Landman point to the core problem of the interpretation of the demonstrative *so*, which has already been mentioned in section (a): what does the demonstrative refer to, and why can it function as a modifier? Siegel's solution is purely surface-oriented. Carlson's account makes use of kinds. Since German *so*, unlike English *such*, has adverbial non-scalar uses in addition to the adnominal ones, adapting Carlson's analysis to German would lead to a proliferation of ontological categories as in Landman's account, and it would rule out a unified account of non-scalar and scalar cases. The present proposal will instead retrieve the modifying property by similarity.

(c) The notion of similarity in Cognitive Science

Similarity is a focal concept in Cognitive Science used to explain a number of **cognitive phenomena** like categorization, learning, reasoning etc. Similarity has been criticized as an explanatory notion by Nelson Goodman as "a pretender, an impostor, a quack" (Goodman 1972, p. 437). He argues that it is meaningless to say that A and B are similar without specifying in which respect, i.e., with respect to which property, they are similar. However, "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." (p. 444).

Goodman's criticism has been countered by numerous authors pointing to a wealth of successful empirical research based on the notion of similarity. Experimental studies revealed **characteristics of similarity**, such as asymmetry (North Korea is rated more similar to Red China than Red China to North Korea), informativeness (when hearing that quaggas are similar to zebras, quaggas are

⁶ These cases improve considerably when taking complete sentences into account: *People in the next room spilled red wine on the floor. Such people are obnoxious.*

assumed to have four legs), and anomaly (statements like *Robins are similar to birds* are judged to be deviant).⁷ Medin and Goldstone (1995) conclude that, in addition to matching and non-matching properties, similarity may depend on factors like directionality and pragmatic inferences. Gentner (1989), moreover, emphasizes the role of relations in determining similarity, distinguishing between analogy, metaphor, and literal similarity.

Sloman and Rips (1998) discuss the question of how compelling similarity as an explanatory construct is, distinguishing a range of positions where, at one end, similarity is viewed as a primitive, fixed by the human cognitive system and, at the other end, it is viewed as parasitic on identity (e.g., identity of shared features). In between are positions viewing similarity as dependent on the set of features relevant in the context, facing the problem of how to determine which features are relevant. Sloman and Rips discuss mental rules as an explanatory alternative to similarity. They conclude "that an adequate model of human thought must take advantage of both the flexibility of **similarity-based** inference and the compositionality and certainty associated with **rule-based** inference" (p. 87).

There are four main approaches to **measure similarity**: the feature-based account, the structural account, the geometrical account, and the transformational account. The feature-based account (Tversky 1977) views similarity as a function of common and distinctive features of the entities under comparison. In the structural account (Gentner and Markman 1997) it is argued that similarity involves the alignment of relational structures. The geometrical account (Gärdenfors 2000) represents similarity in terms of distance in a metrical space. Finally, the transformational account (Hahn et al. 2003) assumes that similarity can be measured by the number and complexity of transformations required to convert one object into the other. Measures of similarity are discussed in Artificial Intelligence using, e.g., topological frameworks and rough sets (Pawlak 1991), and integrating quantitative and qualitative aspects (Lin 2001).

(d) The semantics of gradability and comparison

Approaches to the semantics of gradability and comparison differ, first and foremost, with respect to whether they employ degrees or not. Degree-based analyses interpret gradable adjectives as relations between individuals and degrees, cf. von Stechow (1984) and Bierwisch (1987), or functions from individuals to degrees (Kennedy 1997). In degree-based analyses, the truth conditions of comparative sentences are spelt out with the help of degrees. For example, *Anna is taller than Berta* is true iff the degree **d** such that Anna is **d**-tall exceeds the degree **d'** such that Berta is **d'**-tall. Vague predicate analyses, on the other hand, reject the idea of degrees (e.g. Klein 1980). Gradable adjectives are considered to be of the same type as non-gradable ones, and the truth conditions of comparative sentences are defined with the help of interpretation contexts: *Anna is taller than Berta* is true iff there is a context such that *Anna is tall* is true in this context while *Berta is tall* is false.

Both types of approaches require additional assumptions to explain the vague borderlines of relative adjectives, e.g., *tall*. Degree-based analyses do that by positing a non-crisp threshold (cf. Kennedy 2007), whereas vague predicate analyses allow for truth value gaps (cf., e.g., Sassoon 2008). McNally (2010) accounts for the vagueness of relative adjectives by assuming that their meaning is provided by similarity: an individual is tall if it is sufficiently similar to other tall exemplars. McNally's account is a first step into exploiting the idea of **similarity in the semantics of comparison**.

The analyses of gradability and comparison has mainly been focused on comparative constructions. **Equatives** are commonly considered as a near-identical variant of comparatives indicating a non-strict instead of a strict partial order. While '*A is taller than B*' is true iff A's height is strictly greater than B's height, '*A is as tall as B*' is true iff A's height is greater or equal to B's height.⁸ There are, however, intra-linguistic as well as cross-linguistic differences casting doubt on the near-identity of comparatives and equatives (cf. work package P4).

⁷ Medin and Goldstone (1995) in fact discuss two types of anomaly, '*Robins are similar to birds*.' and '*Robins are similar to questions*.' They argue that the former is anomalous because robins are birds, and the latter is anomalous because similarity judgments presuppose some amount of similarity, which is missing here. Linguists may note that the two types of anomaly correspond to the two types of semantic coordination constraints – semantic difference and the existence of a superordinate concept – proposed by Lang (1984).

⁸ Equatives may also express equality (cf. Atlas 1984), which can, however, be explained by Gricean Maxims.

One of the few detailed analyses of equatives is given in Bierwisch (1987). The analysis is degree-based and built on German examples including a number of valuable linguistic observations, for example, that equative constructions differ from comparative constructions with respect to **norm-relatedness**: While '*Marie ist kleiner als Anna*' ('Mary is smaller than Anna.') does not entail that either of Marie or Anna are small, the sentence '*Marie ist so klein wie Anna*' ('Mary is as small as Anna.') does entail that they are small. Similar issues are discussed for English in Rett (2008).

(e) The grammar of German comparison constructions

The semantic studies of comparison constructions have mainly been focused on English. There are, however, three descriptive studies of German comparison constructions – Hahnemann (1999), Thurmair (2001) and Eggs (2006) – which cover a broad range of data and provide an excellent basis for the proposed project. Hahnemann as well as Thurmair investigate the range of linguistic means to express comparison in German, including equatives, comparatives and superlatives, and also non-scalar constructions. Both give an in-depth **grammatical analysis** of the various syntactic forms of comparison in German and the characteristics of the *als* and the *wie* phrases/clauses introducing the comparison base. Thurmair includes a chapter on the uses of *so* agreeing with Ehlich (1987) in considering the deictic and anaphoric uses as the primary ones, without, however, going into details.

Hahnemann as well as Thurmair propose fine-grained **classifications** of German comparison constructions, including detailed descriptions of the topology and syntactic realization of the components. Thurmair's classification starts from the distinction between equatives, comparatives and superlatives, and comprises four different types of equatives: (i) degree comparison pertaining to adjectives and also gradable nouns and verbs (*so groß wie* 'as tall as'); (ii) kind comparison pertaining to non-scalar nominals (*so ein Auto wie ...* 'such a car as ...'); (iii) comparison of modality pertaining to manner modification (*den Fisch so zerlegen wie ...* 'prepare the fish like ...'); and (iv) comparison of factivity subsuming cases where equatives are used to express additivity (*Anna ist eine Studentin, wie Berta auch*. 'Anna is a student, Berta too.'). Hahnemann's classification takes a different perspective. She starts from the difference between *als* and *wie* in the comparison base and further refines these two categories syntactically. The resulting classification includes some classes for which Hahnemann claims that they are no genuine comparisons, one of them being non-scalar nominal constructions without *so* (*Anna hat ein Auto wie Berta*. 'Anna has such a car as Berta'). Although the notion of "genuine comparison" requires further explication, we will follow Hahnemann in excluding these constructions from the range of equatives considered in the project.

The study by Eggs (2006) takes a broader perspective than the other two, investigating the grammar of *als* and *wie* clauses within and also beyond comparison constructions. Eggs claims that *wie* in general indicates equality, whereas *als* indicates inequality. Although this is questionable, since comparatives with *wie* clauses are more frequent in German than predicted by school grammar (*Anna ist schöner wie Berta*.), the distinction in meaning between *als* and *wie* will be relevant in this project.

(f) More similarity expressions

The expressions *solch*, *genauso*, *ebenso* ('such', 'exactly as', 'same as') are variants of the demonstrative *so* with slight differences in distribution (*solch* can only be combined with nominals), and in usage (*genauso* / *ebenso* can be used anaphorically and in equatives, but not deictically, cf. Thurmair 2001). In addition, there are similarity expressions which are unrelated to the demonstrative *so*, for example, *gleich* and *ähnlich* ('same', 'similar').

The semantics of English *same* (and *different*) has been studied by, e.g., Carlson (1987), Beck (2000) and Lasersohn (2000). There are two studies close to the present proposal: Meier (2009) and Alrenga (2007). Meier proposes a comparative semantics for the expression *resemble* accounting for the fact that it denotes a gradable predicate (*Mary resembles her mother more than her brother resembles her father*). Alrenga (2007) investigates '**similarity comparatives**', i.e. *same*, *different*, *like* (*Barry is the same as his sister*.). He claims that similarity comparatives are basically like scalar comparatives: They measure the degree of (dis)similarity between two individuals, based upon the differences amongst the individuals' locations along various dimensions of comparison. While scalar comparatives relate to only one dimension (*tall*, for example, relates to height), similarity comparatives relate simultaneously to multiple dimensions. This is half way towards the generalized measure phrase approach aimed at in the proposed project (cf. work package P3). However, in Alrenga's account the multiple dimensions

are lumped together in a single artificial dimension R, which is a black box. So there is no way to distinguish between different respects in which individuals are same or different, and there is no way to infer from *A is taller than B* that A is different than B with respect to height, which is a serious drawback.

(g) Contrastive Semantics

Contrastive semantics is a field of growing interest. The basic idea is that languages may be equipped with, or prefer, different methods to achieve the same semantic effect. Quoting the Contrastive Semantics research group at the University of Oslo: "Prospects for progress in the study of natural language meaning are greatly enhanced by paying close attention to the ways in which meaning is conveyed in different languages."⁹

Beck et al. (to appear) present a contrastive semantics study including 14 languages in the area of gradability and comparison, investigating cross-linguistic variation with respect to three parameters: (i) Does the language express comparison by using the positive form and manipulating the context, or by way of degree morphology? (ii) Does it express orderings between individuals, or orderings between individuals and degrees? (iii) Can the degree argument of a gradable predicate be filled overtly? Data representing different comparative forms were elicited by a questionnaire. The results show, for example, that Japanese and English differ with respect to the type of ordering they express: While English licenses orderings between individuals as well as orderings between individuals and degrees, Japanese is restricted to the former type. In a related study, differences in the acquisition of German and English comparatives are accounted for, cf Hohaus & Tiemann (2010).

Summary

The state of the art in the fields relevant to the proposed project leaves open a number of non-trivial questions to be investigated:

1. What does the German demonstrative *so* refer to when used deictically, and how is it possible that a demonstrative acts as a modifier?
2. How to characterize the anaphoric use of the demonstrative *so*? Which types of entities serve as antecedents, and which constraints are imposed by the linguistic material combined with the anaphor?
3. How to make use of the results on similarity in Cognitive Science and Artificial Intelligence in the interpretation of the demonstrative? How to combine truth-conditional semantics with similarity measures and similarity-based classification?
4. Assuming that the expression *so* in German equative comparison is identical to the demonstrative, what is the role of the demonstrative in German equatives? How does it influence the meaning of the equative? How to describe the semantics of non-scalar equatives?
5. What is the difference between the demonstrative *so* and similarity expressions like *genauso*, *gleich*, *ähnlich* ('exactly as', 'same', 'similar')? How do these expressions behave in equative constructions?
6. Do similarity expressions in other European languages (English, Dutch, and Spanish) differ from the German ones? How are scalar and non-scalar equatives expressed in these languages? Are there different strategies to express similarity across languages?

Preliminary work

For each of the above questions, apart from the last two, there exists relevant previous research by the applicant:

ad 1 The meaning and use of **demonstratives** have been the topic of joint work in the Osnabrück Computational Linguistics group for many years, cf. Bosch, Katz, Umbach (2007) and Bosch, Umbach (2007). Ontological questions – which types of referents have to be accounted for in natural language semantics? – have been investigated in the OntoRef project, which was a trilateral cooperation between Osnabrück, Barcelona (Louise McNally) and Lille (Philip Miller) (cf. Umbach 2008a). The use of the demonstrative *so* as an **intensifier** and as a **hedging device** is discussed in Umbach & Ebert (2009), raising the question of why a demonstrative is able to function as an intensifying device and

⁹ <http://www.hf.uio.no/ilos/english/research/groups/contrastive-semantics>

also as a hedging device. The connection between grading and hedging was further investigated in Umbach & Ebert (2010) investigating German specificity marker *gewiss* ('certain') which induces a grading effect when combined with **gradable nouns** (*ein gewisses Risiko*). The intensifier use of the demonstrative *so* when combined with **gradable verbs** (*Er ist so gerast!*) is discussed in Umbach (2011b).

ad 2 Research on the anaphoric use of the demonstrative *so* started in Katz & Umbach (2006), where we considered **anaphoric occurrences of *so*** combined with **verbs**, arguing that the antecedents are predicates instead of manner modifiers. In Umbach (2007a) anaphoric occurrences of *so* combined with **adjectives** are investigated, arguing that that the antecedents are degree modifier but also contextual standards induced by positive forms (*Berta ist [180] groß. Anna ist auch so groß.*). Anticipating the corpus study of the anaphoric uses of *so* scheduled in work package P2, a preparatory study is currently conducted under the guidance of the applicant.

ad 3 In Umbach & Gust (2010) a preliminary approach to make use of the notion of similarity in semantics is presented. The basic idea is to consider nouns as multi-dimensional, in contrast to adjectives, which are one-dimensional, and **generalize the well-established measure function** analysis of gradability from one dimension to many dimensions. This generalization yields functions from individuals to points in **multi-dimensional spaces**, which are familiar in Artificial Intelligence to handle instance-based classification and similarity. While such approaches are often purely quantitative, the one in Umbach & Gust facilitates the combination of qualitative and quantitative information, allowing a variety of different similarity measures.

ad 4 The role of the demonstrative *so* in a transparent analysis of **German equatives** is discussed in Umbach (2007b). The core puzzle is the question of how the degree particle, presuming that it is identical with the demonstrative, relates to the comparison base given by the *wie*-phrase. This work was continued in Umbach (2008b) with a focus on non-scalar equative constructions. It was shown that standard semantics methods are inadequate for their interpretation, which lead to the multi-dimensional approach in Umbach & Gust (2010). The analyses of the **comparative use** and the **additive use** of the focus particle *noch* (*noch schöner, noch ein Schnaps*) in Umbach (2009a) and (2009b) were a spin-off of the research on gradability and comparison providing relevant insights in the problem of norm-relatedness and the concept of additivity.

3. Objectives and work schedule

3.1 Objectives

The key objective of the proposed project is a semantic analysis of expressions of similarity, combining methods from standard truth-conditional semantics with results and techniques from Cognitive Science and Artificial Intelligence. It will be shown that similarity is a fundamental semantic relation, next to identity, manifest in a variety of linguistic expressions.

The center of the project is the German demonstrative *so* in its deictic and anaphoric uses, cf. the examples in (1) and (2) in the previous section. Although the deictic and anaphoric uses are agreed to be the basic uses of the demonstrative, their semantic interpretation is an as yet unresolved issue, since it is unclear what the demonstratives refers to. The analysis developed in the project will solve this problem arguing that the meaning of the demonstrative makes use of similarity instead of identity (cf. work packages P1 and P2).

The similarity relation will be spelt out in a feature-based account generalizing the well-established measure phrase account of gradability to multi-dimensional spaces. This yields a truth-conditional semantics which is able to incorporate different measures of similarity thereby providing a family of similarity relations (cf. P3).

Based on the interpretation of the deictic and anaphoric uses of the demonstrative *so* resulting from work packages P1 and P2, a transparent semantic analysis of German equative comparison will be developed, assuming that the expression *so* in German equatives is an instance of the demonstrative. In addition to scalar equatives, the analysis will include non-scalar ones, which have not been considered in semantics up to now (cf. P4).

Other expressions of similarity, such as *solch*, *genauso*, *gleich*, *ähnlich* ('such', 'exactly as', 'same', 'similar') will be included, comparing their meaning to the interpretation of the demonstrative *so*. The results will create a refined picture of similarity in natural language (cf. P5).

A cross-linguistic study will be conducted, contrasting the analysis of German similarity expressions with data from other European languages (English, Spanish, Dutch), exploring language specific strategies of expressing similarity (cf. P6).

3.2 Work schedule

Work packages – overview (cf. the time schedule in the attachment)

- P1 The deictic use of the demonstrative *so***
- P2 The anaphoric use of the demonstrative *so***
- P3 The relation of similarity in semantics**
- P4 Equative comparison constructions**
- P5 More expressions of similarity**
- P6 Contrastive semantic study of similarity**

Methodology

The research described in the work packages will be based on data resulting from in-depth **corpus studies**. Artificial examples will be used when controlled conditions are required, and also for expository ease. The corpora and search facilities are provided by the Computational Linguistics group at the Institute of Cognitive Science in Osnabrück.¹⁰ The Corpus query tool is CQP (IMS Corpus Workbench). Available corpora are, e.g., the British National Corpus, various German newspaper corpora, and Web corpora in German and English. In addition, there is the multi-lingual EUROPARL corpus containing the debates of the European Parliament with sentence-level alignment (English, German, French and Spanish). Access to multi-lingual corpora will also be granted by the Contrastive Semantics research group at the University of Oslo.¹¹ Moreover, work package P1 includes an **experiment eliciting gestures**, and work package P5 includes an **acceptance rating experiment**.

P1 The deictic use of the demonstrative *so*

The deictic use of the demonstrative *so* will be the starting point of the analysis of similarity expressions in this project. As pointed out in section 2.1, the interpretation of the demonstrative raises two questions, (a) what does the demonstrative refer to?, and (b) how is it possible that a demonstrative acts as a modifier? We argued that a solution based on property theory (Chierchia & Turner 1988), although suggesting itself from a formal point of view, would not be satisfactory, since it doesn't provide an explanation. We will therefore assume, as a working hypothesis, that the semantic ontology includes **individuals** and **events**, but no properties and degrees. Thus the first question of the current work package is: What does the speaker point to in the examples in (1a-c)?

We will follow Nunberg (1993) in assuming that the semantics of a demonstrative comprises (i) a deictic component which refers a contextually salient object, (ii) an interpretation which is the entity contributed to the proposition, and (iii) a relation between the referent and the interpretation, which in the case of run-of-the-mill demonstratives is the identity relation. In the case of the demonstrative *so*, we will assume that it is a **similarity relation**. Consider the adnominal example in (1b), repeated in (5b). According to our hypothesis the referent of the pointing gesture is an individual. The demonstrative *so*, when modifying the phrase *ein Auto*, imposes a constraint such that the referent of this phrase is similar to the individual the speaker points to. Let us call the referent of the demonstration the **comparison base**, and the referent of the modified phrase the **target of comparison**. The demonstrative *so* denotes a relation *sim(comparison-base, target-of-comparison)*. It will be shown in work package P3 how to spell out the similarity relation and endow it with adequate truth-conditions. In the present work package the question will be in focus of how to determine the target of comparison: In the adnominal case in (1b)/(5b), the target corresponds to the interpretation of

¹⁰ cf. <http://cogsci.uni-osnabrueck.de/~CL/resources>

¹¹ cf. <http://www.hf.uio.no/ilos/english/services/omc/>

the NP modified by the demonstrative, i.e. the Anna's car. But what is the target in the adverbial case in (1c)/(5c) and in the ad-adjectival case in (1a)/(5a)? Anna's way of cutting the fish and Anna's height? How to derive these targets compositionally?

- (5) a. (speaker pointing to a person): So groß ist Anna. 'Anna is that tall.'
 b. (speaker pointing to a car): So ein Auto hat Anna. 'Anna has a car like this.'
 c. (speaker pointing to someone dividing a fish):
 So hat Anna den Fisch auch zerlegt. 'Anna cut the fish like that , too.'

The most difficult question when dealing with similarity is the question of **dimensions** – what are the dimensions with respect to which two entities are claimed to be similar? Let us assume, for simplicity, that gradable adjectives are one-dimensional, whereas nouns and verbs are multi-dimensional. In the ad-adjectival case in (5a) the dimension is easy to determine, since it is given by the modified adjective. In the adnominal and the adverbial case, however, the modified noun or verb provide only constraints – in (5b), for example, possible dimensions must be car dimensions. The dimensions relevant for the interpretation must be retrieved from the linguistic and extra-linguistic context, raising the question of how to identify **linguistic constraints** on relevant dimensions, which will be the topic of in-depth corpus search (cf. work package P2).¹²

The deictic use of the demonstrative *so* is characterized by the fact that it requires an accompanying gesture, which is either a pointing gesture or an iconic gesture (cf. Sowa & Wachsmuth 2010). In order to obtain empirical evidence concerning the types of demonstrated entities, a set of gestures accompanying the deictic use of the demonstrative *so* will be compiled. In cooperation with the SFB 673 project "Speech-Gesture Alignment" a sample will be extracted from the **Speech-And-Gesture-Alignment Corpus** (SAGA, cf. Rieser 2010). Moreover, gestures will be elicited in an experimental setting where subjects re-enact dialogs including the demonstrative and are video-recorded. The transcribed recordings will be annotated with a set of distinctive features yielding a classification of possible referents. In addition subjects will be asked to comment their behavior ex-post. The study will be conducted as a joint bachelor thesis of two students.

P2 The anaphoric use of the demonstrative *so*

The anaphoric use of the demonstrative *so* has much in common with the deictic use, for example the problem of determining the relevant dimensions of similarity and the problem of determining the target of comparison. The anaphoric use differs from the deictic one in that antecedents, in contrast to referents of a pointing gesture, are given by linguistic expressions and are thus not restricted to entities one can point to. Although we adhere to the hypothesis that the ontology does not include properties and degrees, the range of individuals qualifying as antecedents will include "abstract objects" like propositions and facts. The challenge for the analysis is to determine **possible antecedents** of the demonstrative.

Consider the adnominal cases in (6). The target of the similarity relation induced by *so* is Anna's car, the comparison base must be retrieved from the preceding discourse. In (6a) (= 2b in section 2.1) the antecedent is obviously the (discourse) referent denoted by *ein Auto mit Heckklappe* – note that the modifier *mit Heckklappe* in itself would not be suited as an antecedent because it cannot be compared to Anna's car. In (6b) the antecedent is obviously the denotation of the NP *ein Cabrio*. Analogous cases led Carlson (1980) to the claim that the antecedents of English *such* are **kinds**, which allow only individual level predicates (cf. section 2.1). Thus one prominent question in the current work package is the question of whether the antecedents of adnominal occurrences of the demonstrative are restricted to kinds.

- (6) a. Berta hat ein Auto mit Heckklappe. Anna hat auch so ein Auto.
 b. Berta hat ein Cabrio. Anna hat auch so ein Auto.
 'Berta has a car with a hatch / a cabrio. Anna has such a car, too.'

¹² One special case of deictic/anaphoric adnominal *so* are occurrences where the noun denotes a dimension, as in *Annas Auto hat auch so eine Farbe* ('Anna's car has such a color, too.').

The examples in (7a,b) show adverbial and ad-adjectival occurrences of the demonstrative. As before, the antecedent has to be an entity that can be compared to the target of the similarity relation, i.e. the event of Anna preparing the fish in (7) and Anna's apartment in (8). Thus the question has to be addressed of what the special characteristics of the antecedents of adverbial and ad-adjectival occurrences are, making use of the findings in Katz & Umbach (2006) and (Umbach 2007a).

- (7) a. Berta hat den Fisch im Wok gegart / gedünstet. Anna hat den Fisch auch so zubereitet.
'Berta cooked the fish in the wok / steamed the fish. Anna prepared the fish in the same way.'
- b. Bertas Wohnung ist 200 qm groß. / ist eine Fabriketage. Annas Wohnung ist auch so groß.
'Berta's apartment is 200 qm. / is a loft. Anna's apartment is that big, too.'

In addition, for all of the anaphoric cases – adnominal, adverbial, ad-adjectival – it has to be tested whether they adhere to the **binding constraints** known for pronouns, cf. (8).

- (8) Jeder Mann ist so schlau wie seine Frau.
'Every man is as clever as his wife.'

The question of how to characterize the antecedents of the demonstrative *so* will be investigated with the help of in-depth **corpus studies**. A preparatory study of adnominal and ad-adjectival cases is currently conducted under the supervision of the applicant. Preliminary results show that the antecedent frequently has to be inferred from previous discourse segments, cf. (9).¹³ Preliminary results also suggest that, in adnominal cases, the majority of the target NPs (the ones modified by the demonstrative) occur in generic sentences, as in the example below. However, before drawing conclusions, the ratio of generic and referential indefinites in general has to be checked.

- (9) Es ist ein großer Verdienst der Mannschaft, dass das Feuer nicht übergegriffen hat und alle angrenzenden Gebäude gehalten werden konnten, [...] erklärte Bürgermeister Friedmann. Wir in der Gemeinde freuen uns, dass wir so eine Feuerwehr haben!
'Thanks to the fire brigade the fire did not spread to the neighbor buildings [...], mayor Friedmann explained. We are glad to have such a fire brigade in our town.'

For ad-adjectival cases, preliminary results show that the adjectives modified by *so* have a **norm-related interpretation**, that is, they entail the positive form. For example, the second sentence in (10) entails that Anna's apartment is big, which is inadequate since the antecedent apartment is clearly small. This is surprising since the corresponding equative doesn't trigger the entailment – how can anaphoricity induce norm-relatedness? This effect vanishes if *so* is substituted by *genauso*, which is another difference between plain *so* and *genauso* (cf. work package P5).

- (10) Bertas Wohnung hat 12 qm. Annas Wohnung ist ??auch so groß./ genauso groß.
'Berta's apartment has 12 qm. Anna's apartment is ??that big, too./ exactly as big (as Berta's).'

The preparatory studies will be extended to include a substantial corpus of anaphoric uses – adnominal, adverbial, and ad-adjectival ones. In the first pass, the data will be manually categorized into anaphoric and non-anaphoric cases. These data will be part-of-speech tagged and semi-automatically annotated with additional syntactic and semantic features, facilitating retrieval of automatic classification criteria.

In cooperation with the Artificial Intelligence group of the Institute of Cognitive Science, two **Machine learning** experiments are planned as jointly supervised bachelor or master theses. The first experiment aims at producing a classifier for anaphoric uses as opposed to non-anaphoric ones. We will test different Machine Learning algorithms and vary the underlying feature set. The main goal is to shed light on properties that are characteristic of anaphoric use of *so*. If the first Machine Learning experiment is successful, the classifier can be used to extract additional examples of anaphoric *so* from the corpora. In the second step, the data sample will be enriched by annotation of possible antecedents and the intended antecedent relation. The enriched sample will constitute the training set in the second Machine Learning experiment aiming at a classifier distinguishing the intended antecedent from other possible ones, which is considerably more sophisticated than the first

¹³ In addition, anaphoric uses of the demonstrative frequently combine with an antecedent in a resultative way – *Sie konnte seine Schweißtropfen sehen. So nah kam sie sonst niemandem.*

experiment. Since it cannot be guaranteed that the Machine Learning experiments are successful, they are only supplementary components of the current project proposal.

P3 The relation of similarity in semantics

In this work package the relation of similarity used in interpreting the deictic and the anaphoric use of the demonstrative *so* will be spelt out such that it can be integrated into a truth-conditional semantics. To begin with, it will be assumed that in the case of ad-adjectival *so* there is one dimension of similarity, i.e. the one named by the adjective, whereas in the case of adnominal and adverbial *so* there are multiple dimensions (cf. P1). The transition from one dimension to many dimensions will require giving up standard predicate logic and make use of multi-dimensional spaces. Moreover, a fully adequate notion of similarity requires giving up classical two-valued logic and use, e.g., a fuzzy notion of truth. Both, multi-dimensional spaces and fuzzy logic are standard in Artificial Intelligence, but not in formal semantics. Since the focus of the proposed project is on the semantics analysis of similarity expressions, and not on a novel formal framework, we will proceed in a conservative manner, refining the similarity relation step by step, such that intermediate steps yield viable solutions for the semantics analysis, even if they are not optimal. This work package will be joint work with Helmar Gust from the Artificial Intelligence group of the Osnabrück Cognitive Science Institute.

The relation of similarity will be spelt out in a feature-based account, which is closer to natural language than, e.g., the geometric account (cf. section 2.1). Similarity is considered as a three-place relation combining two individuals and a (possibly complex) dimension of similarity, which is a function from individuals to suitable values. The first step in characterizing the similarity relation is based on two simplifying assumptions: (a) **one-dimensionality** – there is only one relevant dimension of similarity, and (b) **feature-identity** – similarity is interpreted in the most rigid fashion, that is, two individuals are similar with respect to a dimension if and only if their values in this dimension coincide, cf. (11a), where *x* represents the comparison base, *y* represents the target of comparison and *f* represents the dimension. The demonstrative *so* denotes the property of being similar to the comparison base with respect to dimension *f*, cf. (11b), the comparison base being represented by a free variable *r** to be filled by the referent of the demonstration and the antecedent, respectively. Take the example in (5a) again. Gradable adjectives will be interpreted as measure functions taking individuals to degrees (Kennedy 1999).¹⁴ Measure functions constitute dimensions. Combining *so* with the adjective yields (11c), which is the property of being equal in height to the referent of the demonstration (refDem).¹⁵

- (11) a. $\text{sim}(x, y, f) \text{ iff } f(x) = f(y)$
 b. $[[so]] = \lambda f \lambda x. [f(x) = f(r^*)]$ (where *r** is a free variable)
 c. $[[so \text{ groß}]] = \lambda x. [\text{height}(x) = \text{height}(\text{refDem})]$

The second step in characterizing the similarity relation consists in allowing for more than one dimension, thereby including adnominal and adverbial cases. The idea of the measure function will be generalized from one dimension to multiple dimensions by assuming a function taking individuals to n-tuples of values in the relevant dimensions, that is, taking individuals to points in a multi-dimensional spaces. While in the case of adjectival dimensions like **HEIGHT** the scale is proportional, other dimensions may come with a nominal scale, i.e. an unordered set of values.¹⁶ Imagine, for example, the dimension **EQUIPMENT** including items like *hatch*, *roof spoiler*, *automatic alarm* etc. Similarity for multiple dimensions is spelt out analogous to the one-dimensional case, requiring identity of features, cf. (12a). As before, the demonstrative *so* denotes the property of being similar to the comparison base, but now with respect to multiple dimensions, cf. (12b) (since the selection of relevant features is context-dependent the generalized measure function is represented by a free variable **F***, which is a

¹⁴ This does not mean that degrees are viewed as ontological entities on a par with individuals – they are only auxiliary in description.

¹⁵ Similarity may be implemented such that it mimics the non-strict order which is common for equatives, by using certain closure operations.

¹⁶ In general, scales differ with respect to the relations they provide (identity, partial order) and whether they allow for arithmetic operations.

preliminary solution to be elaborated in the project). Suppose that in the example in (5b) the relevant car dimensions are **EQUIPMENT** and **COLOR**. Combining *so* with the nominal in a compositional way is a non-trivial task due to the pre-determiner position of the demonstrative. Syntactic analyses to be checked in the project are, e.g., in Alexiadou et al. (2007) and Zifonun et al. (1997). Let us for the moment assume that *so* is a regular modifier combining with the noun, like (*ein*) *solches* *N*. Then the result of combining *so* and the noun in (12) is the property of being a car similar to the referent of the demonstration with respect to the relevant car dimensions, (12c).

- (12) a. $\text{sim}(x, y, F)$ iff $f_1(x)=f_1(y) \ \& \ \dots \ \& \ f_n(x)=f_n(y)$ for all components f_i in F
 b. $[[so]] = \lambda x. [F^*(x) = F^*(r^*)]$ (where r^* and F^* are free variables)
 c. $[[so \text{ (ein) } Auto]] = \lambda x. \text{car}(x) \ \& \ [\text{EQUIP}(x) = \text{EQUIP}(\text{refDem}) \ \& \ \text{COL}(x) = \text{COL}(\text{refDem})]$

In the multi-dimensional version above similarity is still implemented by identity of properties, which was the reason why Goodman called similarity a pretender (cf. section 2.1). There are, however, more sophisticated notions of similarity, across dimensions and also within dimensions. Taking the number of common and distinct features into account and adding weights to dimensions will yield a graded notion of **similarity across dimensions** which, when furnished with different thresholds, provides relations of different strength. This is the **third step** in refining the relation of similarity.

The **fourth step** pertains to **similarity within dimensions**. Substituting feature-identity by a less strict relation – let us call it *closeness* to avoid confusion – will allow for an instance based classification procedure (cf. Hahn and Chater 1998). This is what McNally suggested in order to implement the vagueness of relative adjectives (cf. McNally 2010 and section 2.1). One option to implement closeness is the Conceptual Spaces framework in Gärdenfors (2000), which is based on a quantitative similarity measure. In the framework suggested in Umbach and Gust (2010), closeness is implemented in a more general topological space combining qualitative and quantitative aspects. The pros and cons of the two approaches will be evaluated in the project.

P4 Equative comparison constructions

Equatives are the paradigm grammatical construction expressing similarity.¹⁷ German equatives include the expression *so* as a degree modifier. It will be hypothesized in this work package that the degree modifier *so* is identical to the demonstrative *so*, and that the meaning of the demonstrative causes the meaning of the equative comparison. Following this hypothesis, German equatives will be interpreted as expressing similarity – "*Anna ist so groß wie Berta.*" will be analyzed as **sim(Berta, Anna, height)** (but see footnote 16).

Tracing back the meaning of German equatives to the meaning of the demonstrative *so* yields a fully transparent analysis. It might be objected that such an analysis is no longer in parallel to the analysis of comparatives, and that it is language specific. There are, however, a number of up to now neglected data suggesting that there are in fact important differences between equatives and comparatives in German as well as cross-linguistically. For example, **measure phrases** are excluded in German equatives, both as comparison base and as degree modifier, cf. (13).¹⁸ They are licensed in German comparatives, and also in English equatives and comparatives (cf. P6). One other important difference concerns syntax: German equatives, but not comparatives, allow for left-dislocation, cf. (14). Again, English equatives pattern with German comparatives.¹⁹

¹⁷ We will use the following terminology: *Anna* in '*Anna ist so groß wie Berta*' is called the target of comparison, *groß* is the feature of comparison and *Berta* the comparison base. German equative constructions are thus of the form "target of comparison + *so* + feature of comparison + *wie* + comparison base".

¹⁸ cf. Bierwisch (1987). Factorial phrases (*tausend mal*) occur as degree modifiers in both comparatives and equatives. They are, however, not of degree type.

¹⁹ In addition, equatives differ from comparatives with respect to norm-relatedness, cf. Bierwisch (1987). Negative dimensional adjectives do not behave in a norm-related fashion when occurring in comparatives, but they do when occurring in equatives – (a) entails for both Eva and Hans that they are small. Moreover, evaluative adjectives, even positive ones, behave norm-related in equatives – (b) entails that oranges and bananas are tasty. In Umbach (2011a*), the issue of norm-relatedness coming with evaluative adjectives is linked to the discussion about predicates of personal taste and the so-called faultless disagreement.

- (13) a. *Anna ist so groß wie 1.80m. 'Anna is as tall as 1.80m.
 b. *Anna ist 20cm so groß wie Berta. 'Anna is 20 cm as tall as Berta.'
- (14) a. Wie Berta, so groß ist (auch) Anna. lit: 'As Berta, as tall is Anna (too).'
 b. *Als Berta, grösser ist Anna. lit: 'Than Berta, taller is Anna.'

The analysis of German equatives will be to a large extent parallel to the interpretation of the deictic/anaphoric uses of the demonstrative *so*. Equatives differ from deictic/anaphoric uses in the way the comparison base is given. In the deictic/anaphoric uses it is given by a referent or antecedent, whereas in equatives it is given by a *wie*-phrase. As laid out in the previous work package we will assume a measure function analysis of gradable adjectives (Kennedy 1999). The *wie*-phrase will be considered as derived from a clausal structure (*wie Berta groß ist*) by either ellipsis or movement (cf. Lechner 2004), and will be interpreted by applying the measure function denoted by the adjective to an individual. The working hypothesis for the interpretation of (15a) (=3a) is this: The denotation of *so* is the same as in the deictic/anaphoric cases, cf. (16a,b), where similarity is provisionally implemented by feature identity. The *wie*-phrase is interpreted as HEIGHT(berta), which is of type **d** (degree), cf. (16c). The syntactic structure is assumed to be $[[so\ groß] [wie\ Berta]]$ which leaves us with the question of how to combine the two phrases. In the interpretation of the demonstrative suggested for the deictic/anaphoric uses in P1/P2, the comparison base is represented by a free variable, accounting for the fact that it has to be supplied by the context. In the case of the equatives we will try a similar route: introduce the comparison base as a free variable and leave the integration of the *wie*-phrase to a matching mechanism, analogous to the assignment of a value to a free variable, cf. (16d).²⁰ The matching mechanism to be explored in the project is unification, which is a well-established operation in algebra and logic, cf. Goguen (1989).²¹ Comparing the final result in (16d) to the standard interpretation in, e.g., Kennedy (1999), (and neglecting the question of equivalence vs. non-strict order) there is no difference. However, the standard interpretation is derived by viewing the degree particle as directly introducing the order, while the interpretation in (16d) is derived compositionally starting from the meaning of the demonstrative *so*.

- (15) a. Anna ist so groß wie Berta. 'Anna is as tall as Berta.'
 b. Anna hat so ein Auto wie Berta. 'Anna has the same car as Berta.'
- (16) a. $[[so]] = \text{sim}(x, r^*, f) = \lambda f \lambda x. [f(x) = f(r^*)]$ (where r^* is a free variable)
 b. $[[so\ groß]] = \lambda x. [\text{HEIGHT}(x) = \text{HEIGHT}(r^*)]$
 c. $[[wie\ Berta]] = \text{HEIGHT}(\text{berta})$
 d. $[[[so\ groß] [wie\ Berta]]] = \lambda x. [\text{HEIGHT}(x) = \text{HEIGHT}(\text{berta})]$

The equative constructions considered above are scalar, since they involve a gradable expression as a feature of comparison. Non-scalar equatives have, in German, the same surface structure as scalar ones, differing from the latter only in the type of the feature of comparison which is given by a (non-gradable) nominal or verbal expression. As before in the non-scalar deictic/anaphoric cases, the interpretation will be based on a **generalized measure function**, taking individuals to n-tuples of dimensional values.

Consider the example in (15b) (=3b) and imagine, as in P3, that the relevant dimensions are **EQUIPMENT** and **COLOR**. The denotation of *so* is similarity, provisionally implemented by feature identity, (17a). The combination of *so* and *ein Auto* is one of the non-trivial problems to be taken care of in the project, for non-scalar equatives as well as non-scalar deictic/anaphoric cases (cf. P3). We will as before assume the provisional result in (17b). The *wie*-phrase will again be considered as derived by ellipsis or movement (*wie Berta's Auto*) – details have to be examined in the project. The interpretation

-
- (a) *Hans ist so klein wie Eva.* 'Hans is as small as Eva.'
 (b) *Apfelsinen sind so lecker wie Bananen.* 'Oranges are as tasty as bananas.'

²⁰The problem of how to interpret a phrase within the sentence such that it is assigned to a free variable also occurs for dislocation structures, e.g. *Den Onkel, sie hasst ihn* ('The uncle, she hates him.'), cf. Hinterwimmer & Repp (in prep.).

²¹Unification can be performed either per value, unifying the degree resulting from HEIGHT(berta) with the degree variable v^* , or by term, unifying HEIGHT(berta) with $f(r^*)$. If we opt for the latter, as in (16), measure phrases will automatically be excluded.

of the *wie*-phrase consists, as before, in applying the measure function to an individual. Since equipment and color are the relevant dimensions, the generalized measure function has to include two components, which when applied to Berta's car yield a tuple, cf. (17c). Finally, as before we will assume that the denotation of the *wie*-phrase is combined with the denotation of the *so*-phrase by unification. The result is the same as in the deictic/anaphoric non-scalar case, apart from the fact that the comparison base is now given by the *wie*-phrase, cf. (17d)

- (17) a. $[[so]] = \text{sim}(x, r^*, F) = \lambda x. [F^*(x) = F^*(r^*)]$ (where r^* and F^* are free variables)
 b. $[[so \text{ (ein) Auto}]] = \lambda x. \text{car}(x) \ \& \ [\text{EQUIP}(x) = \text{EQUIP}(b's\text{Car}) \ \& \ \text{COL}(x) = \text{COL}(b's\text{Car})]$
 c. $[[wie \text{ Berta's Auto}]] = \langle \text{EQUIP}(b's\text{Car}), \text{COL}(b's\text{Car}) \rangle$.
 d. $[[[so \text{ (ein) Auto}][wie \text{ Berta's Auto}]]] = \lambda x. \text{car}(x) \ \& \ [\text{EQUIP}(x) = \text{EQUIP}(b's\text{Car}) \ \& \ \text{COL}(x) = \text{COL}(b's\text{Car})]$

Hahnemann (1999) as well as Thurmair (2001) consider various types of non-scalar equatives in German which differ with respect to the syntactic category of the comparison base (e.g. NP, PP, clause) and the status of the *wie*-phrase (restrictive / appositive). One particularly interesting finding relates to the presence or absence of the expression *so*. If it is omitted and there is an intonation break, we get an **additive interpretation**. This effect is found in scalar as well as non-scalar examples, cf. (18a,b). However, *so* seems optional in non-scalar examples even if the interpretation is not additive, cf. (18c). According to Hahnemann, these cases do not constitute genuine comparison, which is a claim to be examined in this work package.

- (18) a. Anna ist groß, wie Berta. 'Anna is tall, like Berta.'
 b. Anna hat ein Auto, wie Berta. 'Anna has a car, like Berta.'
 c. Anna hat (so) ein Auto wie Berta. 'Anna has a car like Berta.'

P5 More expressions of similarity

There are more lexical items in German expressing similarity, in addition to the demonstrative *so*, for example *solch* 'such', *genauso* 'exactly as', *gleich* 'same', and *ähnlich* 'similar'. In this work package, these expressions will be compared to the demonstrative *so*, considering distribution, possible uses, and strength of similarity.

Solch can only be combined with nominals, like English *such*, and can be used deictically, anaphorically and in equative constructions.²² Its meaning seems equivalent to *so*. *Genauso* is first of all a modified version of *so* (*genau* meaning *exactly*) and combines with adjectival, nominal and verbal expressions. It can be used anaphorically and in equatives, but it differs from plain *so* in blocking deictic uses, which might be due to accenting patterns. Semantically, *genauso* seems more strict than plain *so*, requiring identical values (in all relevant dimensions), which is obviously due to the meaning of *genau*. Beyond, there are two unexpected differences, the first pertaining to norm-relatedness (entailment of the positive), cf. the example in (10) above, and the second in quantificational contexts – *genauso* excludes negative quantifiers as comparison base, which was already mentioned in Eckardt (2009), cf. (19), and also positive quantifiers in subject position, cf. (20).

- (19) A ist so groß wie niemand anderes / *genauso groß wie niemand anderes.
 'A is as tall / exactly as tall as nobody else.'
 (20) Jeder Mann ist so schlau wie seine Frau / *genauso schlau wie seine Frau.
 'Every man is as clever / exactly as clever as his wife.'

Gleich ('same') and *ähnlich* ('similar') combine with nominal, adjectival, and verbal expressions, but for ad-adjectival and adverbial *gleich* there is only a reciprocal reading, which is not considered in this project. Both *gleich* and *ähnlich* do not allow for deictic uses but only occur anaphorically and in equatives, which is an issue to be examined. Semantically, the hypothesis suggests itself that the former requires identical features for all dimensions constitutive for the type/kind, whereas the latter presupposes that there exist distinct features. There is a significant difference concerning their compatibility with determiners which relate to the strength of similarity: While for *gleich* the definite

²² For cross-linguistic survey of expressions analogous to German *so* / *solch* in English cf. Wood & Vikner (2009)

determiner is required, *ähnlich* as well as *so* and *genauso* combine only with the indefinite determiner, cf. (21a,b). Finally, *ähnlich* is distinct from the other expressions of similarity in being gradable.

(21) a. ?? Anna hat ein gleiches Kleid.

'Anna has a same dress.'

b. Anna hat *das ähnliche Kleid /*auch so das Kleid /*genauso das Kleid.

'Anna has the similar dress / also such the dress / exactly the dress.'

The primary issue in this work package are the differences in **strength of similarity** conveyed by the expressions.²³ It seems that there is a scale, with *gleich* and *genauso* on top, followed by plain *so*, followed by *ähnlich*. This issue will be explored by a corpus study and an experiment collecting online **acceptance ratings**. The corpus study will provide a comprehensive set of occurrences of similarity expressions, which will be part-of-speech tagged and classified along a set of suitable criteria. The experiment will use sentences from the corpus study (plus left and right contexts) manipulated by "cross-splicing", that is, substituting the original similarity expression by one of the others. Eventually, contexts have to be strengthened.

Another issue in this work package relates to characteristics of the similarity relation predicted by findings from Cognitive Psychology, for example asymmetry, anomaly and informativeness, cf. section 2.1. The asymmetry finding will be verified in a corpus study classifying *relata* in similarity relations according to size, importance etc. The properties of anomaly and informativeness are strikingly close to the constraints on coordination in Lang (1984), requiring conjuncts to be (i) semantically distinct and (ii) subsumed by a common superordinated concept. Although these constraints were originally meant to describe conjuncts in coordinated structure, they turned out to be more general (in Umbach 2005* it was argued, e.g., that they apply to focus alternatives). Another corpus study will be conducted to test the hypothesis that Lang's **coordination constraints** apply to *relata* in similarity relations.

Finally, the question arises of why one should use similarity instead of categorical judgments – why say that Anna is as tall as Berta, instead of saying that she is 180cm? One might be tempted to think that the measure phrase is more precise, and the equative is used if the speaker lacks precise information. Consider, however, a scenario where height is not measured in centimeters but in decimeters. In this scenario the measure phrase *18 dm* may be less precise than the equative.²⁴ The difference between the similarity judgment and the categorical one corresponds to the general distinction in Cognitive Psychology between **similarity-based** and **rule-based** (i.e. categorical) classification. The two types of classification were for a long time considered competing models of human cognition. However, Hahn & Chater (1998) convincingly argue that there is no decisive empirical evidence in favour of one or the other mechanism and, moreover, there is evidence that both mechanisms co-exist making use of their complementary strength. From this point of view, it is no surprise to find similarity-based classification alongside categorical classification in natural language.

P6 Contrastive study of similarity (German – English – Spanish – Dutch)

While in the previous work packages the focus was on German data, this one addresses the topic of similarity expressions from a contrastive semantic point of view. The findings for the demonstrative *so* in German will be contrasted with translationally equivalent constructions in English, Spanish and Dutch, exploring other ways to express similarity. This work package will be the core of a PhD thesis.

The parallelism in German (i) between the deictic/anaphoric use of the modal demonstrative *so* and its occurrence in equative constructions, and (ii) between scalar and non-scalar equatives gives rise to the central hypothesis in this project that, first, all of these cases involve the demonstrative *so* expressing similarity, and that, secondly, the interpretation of the equatives is analogous to the deictic and anaphoric cases. English exhibits neither of these parallels, Spanish shows some but not all of them, and Dutch appears fairly close to German. The core issue in this work package is the question of how the four languages differ in expressing the cases in (22a-d). How close are the translational

²³ There is, also, the notorious problem of how to distinguish *das gleiche* and *dasselbe*. The normative rule is such that *das gleiche* denotes type identity while *dasselbe* denotes referential identity. It is unclear, however, whether language use adheres to the rule. To tackle this problem a corpus study carried out in a Bachelor thesis is planned.

²⁴ The argument is, of course, the same with centimeters, but it is intuitively less plausible.

equivalents to the German original? Are there differences in meaning that can be explained in semantic terms? Are there different processes eventually leading to the same semantic results?

(22)	<u>German</u>	<u>English</u>
a. scalar deictic/anaphoric	So groß ist A (auch).	A is that tall / tall like that (too).
b. non-scalar deictic/anaphoric	So ein Auto hat A (auch).	A has such a car / a car like that (too).
c. scalar equative	A ist so groß wie B.	A is as tall as B.
d. non-scalar equative	A hat so ein Auto wie B.	A has such a car as B.
	<u>Dutch</u>	<u>Spanish</u>
a. scalar deictic/anaphoric	A is ook zo groot.	A es así de alta.
b. non-scalar deictic/anaphoric	A heeft ook zo'n auto.	A (también) tiene un coche así .
c. scalar equative	A is zo groot als B.	A es tan alta como B
d. non-scalar equative	A heeft zo'n auto als B.	A tiene un coche como el de B

When considering the examples in (22) it has to be taken into account, first, that each of (a-d) in each of the four languages can, of course, be constructed with the help of other similarity expressions, for example, (22b) in English: *A has the same car as B*. Questions to be examined include, for example, (i) What is the nature of the demonstrative expression in the deictic/anaphoric cases? (ii) How are scalar equatives realized? – Is the degree particle a demonstrative expression? (iii) Are non-scalar equatives constructed in parallel to scalar ones? Methods will include corpus search in multi-lingual corpora and a questionnaire to be developed by the PhD student.

One striking difference across languages as well as between equatives and comparatives consists in the admissibility of measure phrases as a comparison base, cf. (23). Surprisingly, only English allows for measure phrases in equatives.²⁵ In comparatives, measure phrases are allowed in all of the four languages apart from Spanish, which also excludes measure phrases in positive forms (**A es 1,80m alta*. 'A is 180 tall'). The findings in (23) have to be assessed against the background of the results in Beck et al. (to appear). They are additional support for the thesis that there are significant cross-linguistic differences in the way equatives are constructed.

(23)	<u>German</u>	<u>English</u>
a. equative	*A ist so groß wie 1,80m.	A is as tall as 1,80m.
b. comparative	A ist größer als 1,80m.	A is taller than 1,80m.
	<u>Dutch</u>	<u>Spanish</u>
a. equative	*A is zo groot als 1,80m.	*A es tan alta como 1,80m.
b. comparative	A is groter dan 1,80m.	*A es más alta que 1,80m.

The contrastive study in this work package will be complemented by a survey of similarity expressions corresponding to the German ones examined in work package P5, addressing analogous questions (Is the expression a demonstrative, or does it denote a two place relation? Are anaphoric readings available? Can the expression be combined with the definite article? Is it gradable? etc.). The focus will be on how expressions of similarity differ in strength and maybe other characteristics. Bringing together the results will yield a cross-linguistic map of similarity expressions in natural language.

²⁵ According to Rett (2008) such equatives are norm-related entailing that, e.g., 1,80m counts as tall in (23a). Here is a real life example that seems to support her claim: *Most cats have a smell membrane about 14 square centimetres in extent, while that of a dog may be as big as 150 square centimeters.* (DEWAC corpus)

References

Project-related publications

- Bosch, P., Katz, G., Umbach, C. (2007) The Non-subject Bias of German Demonstrative Pronouns. In M. Schwarz-Friesel, M. Consten, M. Knees (eds.) *Anaphors in Text: Cognitive, formal and applied approaches to anaphoric reference*. Benjamins, 145–164.
- Bosch, P. & Umbach, C. (2007) Reference Determination for Demonstrative Pronouns . In D. Bittner & N. Gargarina (eds) *Intersentential Pronominal Reference in Child and Adult Language*, ZAS Papers in Linguistics No 48. Zentrum für Allgemeine Sprachwissenschaft, 39-51.
- Umbach, C. (2009a) Comparatives combined with additive particles: The case of German *noch*. *Proceedings of Sinn & Bedeutung 13*, Stuttgart, 547–562.
- Umbach, C. (2009b) Another additive particle under stress: German additive *noch*. *Proceedings of the Tenth Symposium on Logic and Language*, Hungarian Academy of Sciences, 149 -156.
- Umbach, C. & Ebert, C. (2009) German demonstrative *so*- intensifying and hedging effects. *Sprache und Datenverarbeitung (International Journal for Language Data Processing)* 1-2/2009, 153-168.

Project-related presentations

- Katz, G. & Umbach, C. (2006) *Das auch so*: Anaphoric reference to event predicates and event predicate modifiers (mit Graham Katz), *Sinn und Bedeutung 11*, Barcelona, 09/2006.
- Umbach (2005) Similarity-based Classification in Compositional Semantics. Invited talk. Cognitive Science group, Lund University, September 2005.
- Umbach, C. (2007a) Anaphoric Reference to Gradable Adjectives. *International Conference on Adjectives*, Lille, 09/2007.
- Umbach, C. (2007b) The role of *so* in German equative comparison. *Syntax and Semantics Conference CSSP*, Paris, Oktober 2007.
- Umbach (2008a) Facts as Proofs. Workshop on Reference to Abstract Objects in Natural Language, Universitat Pompeu Fabra, Barcelona, March 2008
- Umbach (2008b) German *wie*-phrases in equative comparison. Invited talk, Universitat Pompeu Fabra, Barcelona, June 2008
- Umbach (2011a) Evaluative predicates and the attitude of '*finden*'. Kolloquium des SFB 673 *Alignment in Communication* der Universität Bielefeld, 01/2011.
- Umbach (2011b) Intensifiers and measure phrases combined with verbs. Workshop *Scalarity in Verb-Based Constructions*, Heinrich-Heine-Universität Düsseldorf, April 2011
- Umbach, C. & Gust, H. (2010) Similarity-based classification in natural language. *10. Jahrestagung der Gesellschaft für Kognitionswissenschaft*. Potsdam October 2010.
- Umbach & Ebert (2010a) Grading by '*gewiss*'. Workshop "Specificity from theoretical and empirical points of view", University of Stuttgart, September 2010.

Other references

- Alexiadou, A., Haegeman, L. & Stavrou, M. (2007) *Noun Phrase in the Generative Perspective*. Berlin: Mouton de Gruyter.
- Alrenga, P. (2007) *Dimensions in the semantics of comparatives*. Dissertation, University of California, Santa Cruz
- Atlas, J. D. (1984) comparative adjectives and adverbials of degree: An introduction to radically radical pragmatics. *Linguistics and Philosophy* 7 347-377.
- Beck, S. (2000) The semantics of *different*: Comparison operator and relational adjective. *Linguistics and Philosophy* 23: 101-139.
- Beck, S., Krasikova, S., Fleischer, D., Gergel, R., Savelsberg, Ch., Vanderelst & J., Villalta, E. (to appear) Crosslinguistic Variation in Comparison Constructions. To appear in *Linguistic Variation Yearbook*.
- 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. N. (1987) *Same and different*. Some consequences for syntax and semantics. *Linguistics and Philosophy* 10: 531-565.
- Castroviejo Miró, E. (2006) A degree-based account of WH-exclamatives in Catalan. *Proceedings of Sinn & Bedeutung* 11, E.Puig-Waldmüller (ed.), Barcelona: Universitat Pompeu-Fabra, pp. 134-149.
- Chierchia, G. & Turner, R. (1988) Semantics and Property Theory. *Linguistics and Philosophy* 11, 261-302.
- Cornish, F. (1992). So Be It: The Discourse–Semantic Roles of So and It. *Journal of Semantics* 9 (2):163-178.
- Dimroth, C., Andorno, C., Benazzo, S., & Verhagen, J. (2010). Given claims about new topics: How Romance and Germanic speakers link changed and maintained information in narrative discourse. *Journal of Pragmatics*, 42(12), 3328-3344.
- Eckardt, R. (2009) „Nothing“ compares to you. Talk at SALT 19, Ohio State University.
- Eggs, F. (2006) Die Grammatik von 'als' und 'wie'. Gunter Narr Verlag, Tübingen.
- 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.
- Gärdenfors, P. (2000) Conceptual Spaces. MIT Press.
- Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7: 155-170.
- Gentner, D., & Markman, A.B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45-56.
- Goguen, J. (1989) *What is Unification?* In M. Nivat, H. Ait-Kaci (eds) *Resolution of Equations in Algebraic Structures, Volume 1: Algebraic Techniques*, Academic Press, 217-261.
- Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (ed) *Problems and Projects*. The Bobbs Merrill Company, Indianapolis and New York, 437–447.
- Hahn, U. & Chater, N. (1998). Similarity and Rules: Distinct? Exhaustive? Empirically Distinguishable? *Cognition*, 65:197–230.
- Hahn, U., Chater, N. & Richardson, L.B. (2003) Similarity as transformation. *Cognition* 87, 1-32.
- Hahnemann, S. (1999) Vergleiche im Vergleich: Zur Syntax und Semantik ausgewählter Vergleichsstrukturen mit "als" und "wie" im Deutschen. Niemeyer, Tübingen.
- von Heusinger, K. (2010) Zwei spezifische indefinite Artikel im Deutschen? Bamberg, Linguistische Werkstatt, Handout, cf. www.ilg.uni-stuttgart.de/vonHeusinger
- Hinterwimmer, S. & Repp, S. (in prep.) Topics(?) in the left periphery.
- Hohaus, V. & Tiemann, S. (2010), "...this much is how much I'm bigger than Joey..." A Corpus Study in the Acquisition of Comparison Constructions", *Proceedings of AcquisiLyon'09*, 90-93.
- 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 Almog, Perry & Wittstein (eds.), *Themes from Kaplan*, Oxford University Press, 481–563.
- Kehler, A. & Ward, G. (1999) On the semantics and pragmatics of 'Identifier so. In K. Turner (ed) *The Semantics/Pragmatics Interface from Different Points of View* (Current Research in the Semantics/Pragmatics Interface Series, Volume I), Amsterdam, Elsevier.
- Kennedy, C. (1999) *Projecting the adjective*. Garland Press, New York.
- Kennedy, C. (2007) Vagueness and grammar: The semantics of relative and absolute gradable adjectives'. *Linguistics and Philosophy*, 30:1-45.

- Kennedy, C. & McNally, L. (2005) Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language* 81, 345-381.
- Klein, E. (1980) A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4:1-45.
- König, E., Stark, D., & Requardt, S. (1990) Adverbien und Partikeln. Ein deutsch-englisches Wörterbuch. Julius Groos Verlag, Heidelberg.
- König, E. (2010) Manner Deixis as a Source of semantic Change and Grammaticalization. Handout, *SWL IV, Dynamique Du Langage*, Lyon.
- 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. Samiiian, (eds.). *Proceedings of the Western Conference in Linguistics (WECOL) 2002*, vol. 11. California State University, Fresno.
- Lang, E. (1984). The semantics of coordination. Studies in Language Companion Series 9. Benjamins, Amsterdam.
- 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.
- Lechner, W. (2004) Ellipsis in Comparatives. Mouton de Gruyter, Berlin.
- Lin, T. Y. (2001) Granulation and Nearest Neighborhood: Rough Set Approach. In Pedrycz W. (ed) *Granular Computing: An Emerging Paradigm*. Physica-Verlag (Springer), 125-142.
- McNally, L. (2011) The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In R. Nouwen, et al. (eds.) *ViC 2009 (Papers from the ESSLLI 2009 Workshop on Vagueness in Communication)*. Lecture Notes in Artificial Intelligence 6517. Berlin: Springer, 151-168.
- Medin, D. L. & Goldstone, R. L. (1995) The predicates of similarity. In Cacciari, C. (ed.) *Similarity in language, thought and perception*. Brepolis. 83-110.
- Meier, C. (2001) Result Clauses. In R. Hastings, B. Kackson, Z. Zvolensky (eds) *SALT XI Proceedings*. 268-285, Ithaca, NY: Cornell University.
- Meier, C. (2009) A Comparative Semantics for Resemblance. In L. Kálmán (ed.) *Proceedings of the 10th Symposium on Logic and Language. LoLa10*, Hungarian Academy of Sciences, 141-148.
- Nouwen (2008) Degrees Under Discussion. Project description, cf. <http://sites.google.com/site/degreesunderdiscussion/home>.
- Nunberg, G. (1993) Indexicality and Deixis. *Linguistics and Philosophy* 16: 1-43.
- Nunberg, G. (2002) Indexical Descriptions and Descriptive Indexicals. In M. Reimer & A. Bezuidenhout (eds) *Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions and Other Related Phenomena*. Oxford University Press.
- Pawlak, Z. (1991). *Rough Sets: Theoretical Aspects of Reasoning About Data*. Dordrecht: Kluwer Academic Publishing.
- Quine, W.V.O. (1960), *Word and Object*, Cambridge MA: MIT Press.
- Rett, Jessica. 2008. "Evaluativity and antonymy." In Gibson, M., Friedman, T. (eds) *Proceedings of SALT XVII*. CLC Publications.
- Rieser, H. (2010). On Factoring out a Gesture Typology from the Bielefeld Speech-And-Gesture-Alignment Corpus (SAGA). In S. Kopp & I. Wachsmuth (eds.), *Proceedings of GW 2009: Gesture in Embodied Communication and Human-Computer Interaction* (LNAI 5934). Berlin/Heidelberg: Springer, 47-60.
- Sassoon, W. G. (2008) Vagueness, Gradability and Typicality, A comprehensive Semantic analysis. Unpublished PhD dissertation, Tel Aviv University.
- Siegel, M. (1994) Such: Binding and the pro-adjective. *Linguistics and Philosophy* 17, 481-497.
- Sloman, St. A. & Rips, L. J. (1998) Similarity as an explanatory construct. *Cognition* 65, 87-101.
- Sowa, T. & Wachsmuth, I. (2010) A computational model for the representation and processing of shape in coverbal iconic gestures. In S. Kopp & I. Wachsmuth (Eds.), *Proceedings of GW 2009: Gesture in Embodied Communication and Human-Computer Interaction* (LNAI 5934), Berlin/Heidelberg: Springer, 132-146.

- von Stechow, A. (1984a) Comparing Semantic Theories of Comparison. *Journal of Semantics* 3, 1-77.
- Tiemann, S., Hohaus, V. & Beck, S. (2010), German and English Comparatives are Different: Evidence from Child Language Acquisition, *Pre-Proceedings of Linguistic Evidence 2010*, 254-258.
- Thurmair, M. (2001) Vergleiche und Vergleichen: Eine Studie zu Form und Funktion der Vergleichsstrukturen im Deutschen. Niemeyer, Tübingen.
- Tversky, A. (1977) Features of similarity. *Psychological Review* 84: 327-352.
- 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.
- Wood, J. & Vikner, S. (2009) Noun phrase structure and movement: a cross-linguistic comparison of *such/ sådan/ solch* and *so/så/so*. *Congres Variation and Change*, University of Amsterdam, handout, cf. www.hum.au.dk/engelsk/engsv/clauses-nominals.
- Zifonun, G., Hoffmann, L. & Strecker, B. (1997) *Grammatik der deutschen Sprache*. De Gruyter, Berlin.