Corpus-based evidence for approximating semantic transparency of complex verbs

Heike Zinsmeister, Eva Smolka

Morphologically complex words possess meanings that range from completely transparent (i. e. similar) to completely opaque (i. e. dissimilar) with respect to the meaning of their base word. German prefix and particle verbs are very productive and frequently used in standard German and are thus a particularly useful means by which to study the effects of relatedness of meaning to the base verb. For example, the particle verbs *auffinden* ('find, locate') and *abfinden* ('compensate, accept') are morphologically derived from the base *finden* ('find'). They thus both share their form with *finden* ('find') though only *auffinden* ('find, locate') shares also its meaning. The linguistic literature (Eisenberg, 2004; Fleischer & Barz, 1992; Olsen, 1996) distinguishes between prefix and particle verbs in that verbal prefixes are inseparable from the base in finite forms (Sie *befindet* sich in X, 'she resides in X') whereas particles are free morphemes and separated from the verb stem in finite forms (Sie *findet* sich mit X *ab*, 'she accepts X'). Besides these morphosyntactic differences, both types of verb derivations may vary with respect to the semantic similarity to the base.

Semantic similarity information is not yet part of lexical databases like CELEX[5] (Baayen, Piepenbrock, & Gulikers, 1995) or dlexDB[6] (Heister et al., 2011). In the following, we argue that

---

[*] Universität Konstanz

[5] WebCelex: `http://celex.mpi.nl/`.

[6] dlexDB: `http://www.dlexdb.de/`.

psycholinguistic experiments would extremely profit from lexical databases that provide information on the semantic similarity between words. First, we present a showcase example of manual data collection for experiments on the lexical representation of complex verbs in German (cf. Smolka, Komlósi, & Rösler, 2009; Smolka, Preller, & Eulitz, 2011). Second, we outline a pilot study that extracts semantic similarity between complex verbs and their base verbs from a corpus.

The behavioral experiments in German question the hypothesis based on findings in English (cf. Feldman, Barac-Cikoja, & Kostić, 2002; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Rastle, Davis, Marslen-Wilson, & Tyler, 2000) that meaning compositionality determines the lexical representation of complex words. That is, if complex words are semantically transparent, they are lexically represented via their base, otherwise they are assumed to be represented as whole words. In contrast to the findings in English, though, several experiments have shown that German complex verbs are represented via their base regardless of meaning compositionality. That is, both semantically transparent and opaque complex verbs produced equivalent priming effects. The base *binden* ('bind') was primed to the same extent by transparent verbs like *zubinden* ('bind together') as it was by opaque verbs like *entbinden* ('deliver'). Moreover, the priming by morphologically derived verbs was significantly stronger than that by purely meaning related verbs like *zuschnüren* ('bind together').

These findings indicate that the lexical representation of complex words in German refers to the base regardless of meaning. Lexical representation in German thus differs from that in other Indo-European languages.

Providing valid evidence requires to strictly control the stimulus materials for usage-based variables like lemma frequencies on the one hand and the assessment of semantic similarity on the other hand. While many distributional variables are provided by lexical databases, meaning similarity needs to be assessed by laborious means: For each experiment (cf. Smolka et al., 2009, 2011), the collection of the stimuli started out by choosing between 66 to 80 from the roughly 1400 monomorphemic German verbs in CELEX, and by manually listing the possible prefixed derivations for each of them and ascertaining whether they convey a similar meaning as the base verb. The latter decision was operationalised according to Hay's (2001) proposition that complex words are semantically transparent, if the base is used in a lexical paraphrase of the complex verb. In addition, semantic association tests were conducted to establish the meaning relatedness between primes and targets for all prime conditions. For each of the candidate words, two semantically transparent derivations, two semantically opaque derivations, and various control words were distributed across lists. In total, up to 640 prime-target pairs were tested, and between 80 and 120 participants who did not participate in the experiments proper rated the meaning relatedness between the verbs of each pair on a 7-point scale from completely unrelated (1) to highly related (7). According to different thresholds for the different prime conditions (e.g., mean ratings > 4 for semantically related conditions), base verbs and their primes were included in the critical set. One finding was that these judgments often diverted from the transparency definitions based on the lexicon paraphrases, so that often more than one association test needed to be conducted. All in all, the creation of the stimulus materials is extremely laborious and the overall duration of the stimulus preparation takes about 2-3 months. Supporting lexical databases with corpus-based similarity information is thus extremely warranted.

In computational linguistics, semantic similarity is often determined by distributional similarity[7]: two words are semantically similar, if they occur in similar contexts, that is, if they co-occur with a similar set of words (e.g., Landauer & Dumais, 1997) or have similar selectional preferences (e.g., Erk & Pado, 2010). Earlier distribution-based work on verb compositionality include Baldwin, Bannard, Tanaka, and Widdows (2003) and McCarthy, Keller, and Carroll (2003) on English[8] and Schulte im Walde (2005) on German. Kühner and Schulte im Walde (2010) studied the transparency of German particle verbs with respect to their base verbs. They presented two soft-clustering algorithms on the basis of selectional preferences and evaluated the resulting clusters against a gold standard of human association judgments. They obtained best correlations between clustering and human association scores with clustering that was based on argument fillers like subject fillers (higher correlation but lower coverage) and prepositional object fillers (lower correlation but better coverage). Object fillers proved to be less useful than the subject fillers, which was explained by the fact that German complex verbs often differ in their subcategorization frames from that of the base verbs.

(5.1) Sie lächelte vielsagend.
'She smiled tellingly'

(5.2) *Sie lächelte [$_{NPacc}$ ihre Mutter] vielsagend.
'She smiled her mother tellingly.'

---

[7] See Baroni and Lenci (2010) for a recent survey on the state-of-the-art of distributional semantics. They also introduce Distributional Memory DM, `http://clic.cimec.unitn.it/dm/`, a resource that models English as a set of weighted *word-link-word* tuples arranged into a multi-dimensional data structure, whereby *link* is substituted by 'syntagmatic links', for example, a verb that links two nouns. DM was trained on 2.83 billion tokens pos-tagged and dependency parsed text. It is thus a general, task-independent resource, which may serve as a basis for further computing of similarity scores – as would be needed for psycholinguistic databases. A similar but smaller resource for German is provided by the DWDS *Wortprofile* 'word profiles', (cf. Geyken, Didakowski, & Siebert, 2009).

[8] Wulff (2010) presents a recent study on the compositionality of English V NP-idioms.

(5.3) Sie lächelte [$_{NPacc}$ ihre Mutter] vielsagend an.
'She smiled at her mother tellingly.'

The Examples (5.1)-(5.3) (modified from Kühner & Schulte im Walde, 2010) show that *lächeln* ('smile') is an intransitive verb, whereas the derived particle verb *anlächeln* ('smile at') is transitive and subcategorizes for an accusative object. The object filler *Mutter* 'mother' of *anlächeln* would, therefore, not model the meaning similarity between the two verbs. Different from the subcategorization, both verbs can be modified by the adverb *vielsagend* ('tellingly').

In the present study, we explore whether modifiers of the verb can be used to adequately model the semantic transparency of complex verbs. This was motivated by the fact that the semantics of a predicate strongly determines the range of its modifiers, since the modifiers relate to temporal, spatial or other properties of the verbs. We thus expected that complex verbs co-occur with a set of modifiers that is similar to that of their base verb, if they are semantically transparent. In contrast, the set of modifiers will differ from that of the base, if the complex verb possesses a different meaning than its base.

In a first pilot study, we investigated the distributional behavior of 45 base verbs and their derivatives that have been previously used in the association tests and priming experiments of Smolka et al. (2009). The distributional data are extracted from a parsed subcorpus (about 60 million tokens) of the SDeWaC corpus (cf. Faaß, Heid, & Schmid, 2010, based on the DeWaC web corpus Baroni, Bernardini, Ferraresi, & Zanchetta., 2009). After parsing the corpus by a dependency parser (Bohnet, 2010), we collected frequencies of verb lemmas and how often they co-occurred with individual fillers of the dependents' heads. The verb lemma frequencies sum the frequencies of a verb in all its inflectional forms, non-finite and finite, the latter including occurrences of the verbs with stranded separated particles. Complex verbs contribute their own lemma and do not add to the frequency of the lemma of their base verb. In the sentence *Ihre Nation sieht Tiere als Freunde an*, see Table 5.1, the particle verb *ansehen* ('consider') is separated into the finite verb *sieht* and the particle *an*, which is related to its head *sieht* by means of the dependency relation SVP ('separated verb particle'). *Nation* is the head of the subject (SB), *Tiere* the head of the direct object (OA), and the preposition *als* the head of the modifier (MO).[9] The part-of-speech tags in the column 'pos' belong to the STTS tagset for part of speech tagging (cf. Schiller, Teufel, Stöckert, & Thielen, 1999).

Table 5.1: Dependency analysis of *Ihre Nation sieht Tiere als Freunde an* ('Their nation considers animals as friends'). Particle verb: *ansehen* ('consider')

ID	word	gloss	lemma	pos	regent	rel
1	Ihre	'their'	ihr	PPOSAT	2	NK
2	Nation	'nation'	Nation	NN	3	SB
3	**sieht**	'considers'	sehen	VVFIN	0	ROOT
4	Tiere	'animals'	Tier	NN	3	OA
5	als	'as'	als	APPR	3	MO
6	Freunde	'friends'	Freund	NN	5	NK
7	**an**	AN	an	PTKVZ	3	**SVP**

The parser was reported to achieve 88.06 % labeled attachment score for German, outperforming the best systems of a previous shared task in dependency parsing (cf. Bohnet, 2010). To judge the reliability of our corpus, we performed two small-scale evaluations, first, regarding verbal dependents in general and, second, regarding separated particle verbs in particular.

With regard to the analysis of verbal dependents, the 30 test sentences had an average length of 30.9 words without punctuation (min: 16 words, max: 39 words). The evaluation was restricted to the relations that are exploited in the present study: SB, OA, and MO–the latter with a further specification for certain heads like preposition (APPR), adjective (ADJD) or adverb (ADV). Table 5.2 summarizes the evaluations of the verbal dependents as well as

---

[9] More precisely, *als Freunde* is a verbal argument that predicates over the direct object. However, the parser was trained on the TIGER treebank, which provides a modifier analysis in cases like this (cf. Albert et al., 2003, p. 86).

those for all dependents together (dep$_{all}$). ‘Gold freq’ lists the actual frequencies in the test corpus.

Table 5.2: Labeled attachment scores based on 30 test sentences

Relation	gold freq	precision	recall	f-score
dep$_{all}$	205	0.83	0.96	0.89
SB	65	0.86	0.94	0.90
OA	35	0.77	0.86	0.82
MO	105	0.83	1	0.91
MO$_{APPR}$	50	0.83	1	0.91
MO$_{ADJD}$	12	0.92	1	0.96
MO$_{ADV}$	35	0.81	1	0.90

For the evaluation of separated particle verbs, we compiled a small test corpus of 165 separated occurrences of six different particle verbs (with test sets of 30–or 15–occurrences of each verb). The verbs for this task covered different lemma frequencies and different proportions of separated vs. non-separated occurrences, see Table 5.3. The parser achieved an average precision of 97.58 % in identifying the particle verbs that occurred in the form of finite verb plus separated particle.

Table 5.3: Sample precision scores for the parsing of separated particle verbs

Verb	lemma freq	separated	test set	precision
ansehen (‘consider, look at’)	7313	0.20	30	0.97
aufrufen (‘access, call up’)	2275	0.46	30	1
vormachen (‘demonstrate, fool’)	280	0.28	30	0.90
zurückbleiben (‘stay behind’)	709	0.67	30	1
zusammenkneifen (‘squint, punch’)	19	0.79	15	1
zuziehen (‘contract, consult’)	392	0.54	30	1

We concluded from these evaluations that (i) immediate dependents of a verb are identified correctly in most cases (cf. the high recall scores in Table 5.2), even though (ii) the parser tends to link additional items to the verbs (cf. the lower precision scores), which may blur the context sets for defining semantic similarity. Most importantly, most of the particle verbs that occurred as finite verb and separated particle had been identified in a reliable way. The last finding is encouraging, since it indicates that all inflectional forms of particle verbs, both separated and continuous strings, can be used to extract context evidence.

The parsed corpus suggested that roughly one third of the text occurrences of each particle verb is of the form finite verb plus separated particle.[10] Excluding these types would weaken the empirical basis for the investigation of particle verbs in comparison to other verb types.

For the present study we thus used all data, including those of separated particle verbs. We created context vectors for 279 verbs, which comprised 45 base verbs and about five additional verbs per base verb-derived and non-derived, semantically related and non-related. The context vectors saved the co-occurrence frequencies of the fillers of the dependents' heads, namely, of subjects (SB), direct objects (OA), adjectival modifiers (MO_{ADJD}), adverbial modifiers (MO_{ADV}), and prepositional modifiers (MO_{APPR}). A context vector for the example in Table 5.1 is sketched in Example (5.4). With each subsequent occurrence of the verb the frequencies of co-occurring fillers are increased and additional fillers are added, such as, fillers of adjectival or adverbial modifiers that are not provided in the example sentence.

(5.4) `ansehen SB: Nation 1 OA: Tiere 1` $\mathtt{MO}_{ADJD}$`:` $\mathtt{MO}_{ADV}$`:`
`MO`$_{APPR}$`: als 1`

The verbs were instantiated by probability distributions over their dependents. Following Schulte im Walde (2005), we computed the

---

[10] The mean proportion of separated occurrences calculated for 73 particle verbs with a lemma frequency larger than 20 was 31 % (1st Quartile=0.21, 3rd Quartile=0.41, Maximum=0.67).

distance d between two verbs v_1 and v_2 with a smoothed variant of the Kullback-Leibler divergence D (cf. Equation 5.5), by skew divergence D^s (Lee, 2001, cf. Equation 5.6). Skew divergence D^s the distance between the probability distribution p of verb v_1 (in our case the base verbs) for co-occurring with its dependents' fillers i to the probability distribution q of verb v_2 (in our case the derived verbs) for co-occurring with all of the dependents' fillers i of verb v_1. Skew divergence smoothes for zero probabilities of q_i, if verb v_2 never co-occurred with a certain dependent filler of verb v_1 by replacing a fraction of q_i with a fraction of the probability mass of p_i.

(5.5) $d(v_1, v_2) = D(p||q) = \sum_i p_i * log \frac{p_i}{q_i}$

(5.6) $d(v_1, v_2) = D(p||w * q + (1 - w) * p)$

The lower the skew divergence score D^s, the smaller the distributional distance between the two verbs. Hence, only if the two verbs occurred with identical dependents throughout the corpus, the score will arrive at zero. Put it differently, the lower the skew divergence score the greater the distributional similarity between the two verbs.

Table 5.4 provides an example for the comparison of skew divergence scores and human association scores: the scores for the base verb *bleiben* ('stay'), three derived verbs (*zurückbleiben* 'stay behind', *aufbleiben* 'stay up', *unterbleiben* 'stop') and two other verbs (*bestehen* 'persist', *senken* 'cut, reduce'). The second column *type* indicates whether the verbs are semantically related (+S) or form related (+F) with the base *bleiben*. The third column lists the lemma frequencies of the corpus. The column H provides the medians of the human association scores and R_H the corresponding similarity ranking of the verbs.[11] $D^s_{dep\_all}$ provides the skew divergence scores calculated on the basis of the set of all dependents and R_{dep_all} the corresponding ranking. Finally, for reasons of comparison, the last two columns show the skew divergence scores D^s_{MO} calculated on the set of modifiers only and its corresponding ranking R_{MO}.

---

[11] Identical scores are assigned the same normalized rank.

Table 5.4: Comparison of distributional similarity and human association measures

Verb	type	freq	H	R_H	$D^s_{dep\_all}$	R_{dep_all}	D^s_{MO}	R_{MO}
bleiben ('stay')		36590						
zurückbleiben ('stay behind')	+S+F	706	5.5	1^{st}	1.27	2^{nd}	0.32	1^{st}
bestehen ('persist')	+S-F	21313	5	2^{nd}	1.19	1^{st}	0.41	2^{nd}
aufbleiben ('stay up')	?S+F	14	3.5	3^{rd}	3.12	5^{th}	1.56	5^{th}
unterbleiben ('stop')	-S+F	300	1	4.5^{th}	2.11	4^{th}	0.81	4^{th}
senken ('cut, reduce')	-S-F	2140	1	4.5^{th}	1.76	3^{rd}	0.63	3^{rd}

The human association scores H in Table 5.4 cluster the verbs into three groups: the semantically related ones (scores ≥ 5), the unrelated ones (score = 1) and the intermediate one (score = 3.5). These clusters are also reflected in the corpus-based rankings R_{dep_all} and R_{MO}: the semantically related verbs rank on 1^{st} and 2^{nd} rank and the unrelated ones on 3^{rd} and 4^{th}, while the verb *aufbleiben* ('stay up') is not well classified by the corpus-based scores probably due to its sparseness in the corpus (frequency = 14).

The distribution-based distances were then compared with the human association scores using the Spearman rank-order correlation coefficient (Siegel & Castellan, 1988; cf. McCarthy et al., 2003 and Kühner & Schulte im Walde, 2010). We expected a negative correlation due to the inverted scales of divergence scores and human association scores. We compared the scores of all 236 verb pairs (derived from the 45 base verbs). As provided in Table 5.5, the best correlation was achieved when the divergence scores were calculated on the basis of all dependents $D^s_{dep\_all}$ (one-sided Spearman's rank correlation test: ρ = -0.22, S = 2634507, $p < 0.001$). Even though the correlation is significant it is rather low. Therefore we looked into the data in more detail.

Table 5.5: Correlation ρ between skew divergence scores D^s and human association scores

D^s type	ρ	p-value
$D^s_{dep\_all}$	-0.22	$p < 0.001$
D^s_{SB}	-0.13	$p < 0.05$
D^s_{OA}	-0.09	n.s.
D^s_{MO}	-0.18	$p < 0.01$
$D^s_{MO\_APPR}$	-0.16	$p < 0.05$
$D^s_{MO\_ADV}$	-0.09	n.s.
$D^s_{MO\_ADJD}$	-0.10	n.s.

Comparing subsets of the verb pairs (i.e., verb pairs v_{high} with high human scores H of 6 or 7 versus verb pairs v_{low} with low human scores H of 1 or 1.5) showed that the divergence scores $D^s_{dep\_all}$ of verb pairs v_{high} have a significantly higher distributional similarity $D^s_{dep\_all}$ than pairs of the group v_{low} (according to the one-sided Wilcoxon rank sum test with continuity correction, W = 141930, $p < 0.05$).[12] This indicates that the two measures, H versus D^s, cluster the data in a similar way, at least with regard to the extreme ends of the scales.

Furthermore, we tested the correlation between H and D^s with respect to the scores of individual base verbs. 13 out of the 45 base verbs (29 %) showed at least one significant variant of the skew divergence score D^s. The correlation coefficient *rho* was relatively high (ranging between -0.74 and -0.94, $p < 0.05$), indicating high correlation. However, there was no clear pattern with respect to the different types of fillers.

One explanation for the low overall correlation between D^s and H may be that the data for calculating the corpus-based evidence are too sparse. In this pilot study, the calculations have been based on the lemmas of the dependents' heads. It is possible that better results would be achieved, if lemmas were generalized to more

---

[12] We used this non-parametric test since the data were not normally distributed. An alternative approach would have been to test the log-transformed values as suggested by Gemma Boleda p.c.

abstract lexical concepts. Furthermore, more specific sets of the modifier relation (MO) and its subclasses may model the verbs in a more meaningful way. At last, future studies will need to address the problem of polysemy of the verbs both in corpus-based evidence and human association scores.

To summarize, using dependency parses allows us to exploit verbs with separated verb particles. This enlarges the empirical basis and boosts the frequency counts for verbs and their dependents that are lost in other approaches.

Our results suggest that the meaning similarity between verbs can be well modeled by the distributional similarity based on modifiers in addition to that of arguments.

Acknowledgements

We thank Bernd Bohnet, who helped us with parsing the corpus data. We also thank the audience of the Workshop on lexical resources in psycholinguistic research, Berlin, March 2011, in particular Gemma Boleda, Alexander Geyken and Amir Zeldes for helpful comments and advice. The project was partly supported by the Europäischer Sozialfonds in Baden-Württemberg.

Contact: Heike Zinsmeister
`<heike.zinsmeister@uni-konstanz.de>`

References

Albert, S., Anderssen, J., Bader, R., Becker, S., Bracht, T., Brants, S., et al. (2003, Juli). TIGER Annotationsschema [Computer software manual].

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The celex lexical database (cd-rom)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baldwin, T., Bannard, C., Tanaka, T., & Widdows, D. (2003, July). An empirical model of multiword expression decomposability.

In *Proceedings of the acl 2003 workshop on multiword expressions: Analysis, acquisition and treatment* (pp. 89–96). Sapporo, Japan.

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta., E. (2009). The WaCky Wide Web: A collection of very large linguistically processed Web-crawled corpora. *Journal of Language Resources and Evaluation*.

Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, *36*(4), 673–721.

Bohnet, B. (2010, August). Top accuracy and fast dependency parsing is not a contradiction. In *Proceedings of the 23rd international conference on computational linguistics (coling 2010)* (pp. 89–97). Beijing, China.

Eisenberg, P. (2004). *Grundriß der deutschen Grammatik: Das Wort* (2nd ed., Vol. 1). Stuttgart-Weimar: J. B. Metzler.

Erk, K., & Pado, S. (2010, July). Exemplar-based models for word meaning in context. In *Proceedings of the acl 2010 conference short papers* (pp. 92–97). Uppsala, Sweden.

Faaß, G., Heid, U., & Schmid, H. (2010). Design and application of a Gold Standard for morphological analysis: SMOR as an example of morphological evaluation. In *Proceedings of the seventh international conference on language resources and evaluation (lrec)* (pp. 803–810). Valetta, Malta.

Feldman, L., Barac-Cikoja, D., & Kostić, A. (2002). Semantic aspects of morphological processing: Transparency effects in Serbian. *Memory and Cognition*, *30*, 629–636.

Fleischer, W., & Barz, I. (1992). *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Max Niemeyer Verlag.

Geyken, A., Didakowski, J., & Siebert, A. (2009). Generation of word profiles for large German corpora. In Y. Kawaguchi, M. Minegishi, & J. Durand (Eds.), *Corpus analysis and variation in linguistics* (pp. 141–157). Tokyo: Benjamins.

Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics: An Interdisciplinary Journal of the Language Sciences*, *39*(6), 1041–1070.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*, 10–20.

Kühner, N., & Schulte im Walde, S. (2010). Determining the Degree of Compositionality of German Particle Verbs by Clustering Approaches. In *Proceedings of the 10th conference on natural language processing* (pp. 47–56). Saarbrücken, Germany.

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: the Latent Semantic Analysis Theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lee, L. (2001). On the effectiveness of the skew divergence for statistical language analysis. In *Artificial intelligence and statistics 2001* (pp. 65–72).

Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33.

McCarthy, D., Keller, B., & Carroll, J. (2003, July). Detecting a continuum of compositionality in phrasal verbs. In *Proceedings of the acl 2003 workshop on multiword expressions: Analysis, acquisition and treatment* (pp. 73–80). Sapporo, Japan.

Olsen, S. (1996). Partikelverben im Deutsch-Englischen Vergleich. In E. Lang & G. Zifonun (Eds.), *Deutsch – typologisch* (pp. 261–288). Berlin: De Gruyter.

Rastle, K., Davis, M., Marslen-Wilson, W., & Tyler, L. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*, 507–537.

Schiller, A., Teufel, S., Stöckert, C., & Thielen, C. (1999). *Guidelines für das Tagging deutscher Textcorpora mit STTS* (Tech. Rep.). Stuttgart: Institut für maschinelle Sprachverarbeitung.

Schulte im Walde, S. (2005). Exploring Features to Identify Semantic Nearest Neighbours: A Case Study on German Particle Verbs. In *Proceedings of the international conference on recent advances in* *natural language processing* (pp. 608–614). Borovets, Bulgaria.

Siegel, S., & Castellan, N. (1988). *Non-parametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Smolka, E., Komlósi, S., & Rösler, F. (2009). When semantics means less than morphology: Processing German prefixed verbs. *Language and Cognitive Processes*, *24*(3), 337–375.

Smolka, E., Preller, K., & Eulitz, C. (2011). *The lexical representation of complex verbs in German: Morphological structure overrides semantic compositionality.* (Manuscript submitted for publication)

Wulff, S. (2010). Marrying cognitive-linguistic theory and corpus-based methods: On the compositionality of English V NP-idioms. In D. Glynn & K. Fischer (Eds.), *Corpus-driven cognitive semantics.* Berlin/New York: Mouton.