Good and bad opposites

Using textual and experimental techniques to measure antonym canonicity

Carita Paradis, Caroline Willners and Steven Jones
Växjö University, Sweden / Lund University, Sweden / The University of Manchester, UK

The goal of this paper is to combine corpus methodology with experimental methods to gain insights into the nature of antonymy as a lexico-semantic relation and the degree of antonymic canonicity of word pairs in language and in memory. Two approaches to antonymy in language are contrasted, the *lexical categorical* model and the *cognitive prototype* model. The results of the investigation support the latter model and show that different pairings have different levels of lexico-semantic affinity. At this general level of categorization, empirical methods converge; however, since they measure slightly different aspect of lexico-semantic opposability and affinity, and since the techniques of investigation are different in nature, we obtain slightly conflicting results at the more specific levels. We conclude that some antonym pairs can be diagnosed as "canonical" on the strength of three indicators: textual co-occurrence, individual judgement about "goodness" of opposition, and elicitation evidence.

Keywords: adjective, antonym, contrast, synonym, gradable, prototype, conventionalization, lexico-semantic relation

It has long been assumed in the linguistics literature that contrast is fundamental to human thinking and that antonymy as a lexico-semantic relation plays an important role in organizing and constraining languages' vocabularies (Cruse, 1986; Fellbaum, 1998; Lyons, 1977, M. L. Murphy, 2003; Willners, 2001).[1] While corpus methodologies and experimental techniques have been used to investigate antonymy, little has been done to combine the insights available from these methods.[2] The purpose of this paper is to fill this gap and shed new light on lexico-semantic relations in language and memory.

This article centres on the notion of antonym canonicity. Canonicity is the extent to which antonyms are both semantically related and conventionalized as

The Mental Lexicon 4:3 (2009), **380–429**. DOI 10.1075/ml.4.3.04par
ISSN 1871–1340 / E-ISSN 1871–1375 © John Benjamins Publishing Company

pairs in language (M. L. Murphy, 2003, p. 31). A high degree of canonicity means a high degree of lexico-semantic entrenchment in memory and conventionalization in text and discourse, and a low degree of canonicity means weak or no entrenchment and conventionalization of antonym couplings. The lexical aspect of canonicity concerns *which* words pairs are located *where* on a scale from good to bad antonyms and the semantic part focuses on *why* some pairs might be considered better oppositions than others. This study measures which adjectives form part of strongly conventionalized antonymic relations and which adjectives have no strong candidate for this relationship. For instance, speakers may readily identify *fast* as the antonym of *slow*, but may be less confident in assigning an antonym to, say, *rapid* or *dull*. When asked to make judgements about how good a pair of adjectives are as opposites, speakers are likely to regard *slow – fast* as a good example of a pair of strongly antonymic adjectives, while *slow – quick* and *slow – rapid* may be perceived as less good pairings, and *fast – dull* a less good pairing than *slow – quick* and *slow – rapid*. All these pairs in turn will be better examples of antonymy than pairs such as *slow – black* or synonyms such as *slow – dull*.

Our hypothesis is that there is a limited core of highly opposable couplings that are strongly entrenched as pairs in memory and conventionalized as pairs in text and discourse, while all other couplings form a scale from more to less strongly related. This hypothesis is consistent with prototype categorization and will be referred to as the *cognitive prototype approach* (cf. Cruse, 1994). Our approach challenges the *lexical categorical approach* to antonymy, which argues that a strict contrast exists between two distinct types of *direct* (i.e., lexical) and *indirect* antonyms, and that such a dichotomy is context insensitive as assumed in some of the literature (e.g., Princeton WordNet, Gross & Miller, 1990). Unlike Gross and Miller's categorical approach, which is a lexical associative model, we argue that antonymy has conceptual basis and meanings are negotiated in the contexts where they occur. However, in addition, there is a small set of adjectives that have special status in that they also seem to be subject to lexical recognition by speakers. For instance, it is perfectly natural to ask any native speaker including small children what the opposite of *good* is and receive an instantaneous response, while the opposite of, say, *grim* or *calm* would create uncertainty and require some consideration on the part of the addressee. Similarly, asking for a word that means the same as *good* does not give rise to an immediate response and the question is not easily answered by small children.

The study is situated within the broad Cognitive Linguistics framework (Croft & Cruse, 2004; Langacker, 1987; Talmy, 2000), in which meanings are mental entities and arise through context-driven conceptual combinations. Words activate concepts; lexical meaning is the relation between words and the parts profiled in meaning-making. There is no way we can pin down the meaning of words out

of context. If we do not have a context, we automatically construct a context. Lexical meanings are constrained by encyclopaedic knowledge, conventionalized couplings between words and concepts, conventional modes of thought in different contexts and situational frames. Words do not *have* meanings as such; rather, meanings are evoked and constantly negotiated by speakers and addressees at the time of use (Cruse, 2002; Paradis, 2003, 2005). They function as triggers of construals of conceptual structures and cues for innumerable inferences in communication (G. L. Murphy 2002, p. 440; Verhagen, 2005, p. 22). Cognitive Linguistics is a usage-based theory in the sense that language structure emerges from language use (e.g., Langacker, 1991; Tomasello, 2003). Some linguistic sequences are neurologically entrenched in our minds through co-occurrence of use, while others are loosely or not at all connected because of a weak collocational link in language or because they are occasional.

In mental lexicon research, an important distinction is made between stored knowledge (representations) and computation (cognitive processing and reasoning) (Libben & Jarema, 2002). The two approaches which are contrasted in this article represent two different views on the role of representations and reasoning. The categorical approach relies heavily on stored static lexical associations. Relations in that approach are primitives, and meanings are not substantial but derived from the relations. Within the cognitive, continuum approach, on the other hand, meanings are conceptual in nature and relations, such as antonymy, are construal configurations and produced by general cognitive processes, such as attention, *Gestalt* and comparison (Paradis, 2005). Construals form the dynamic part of the model. They operate on the conceptual pre-meanings in order to shape the final profiling when they are being used in communication (for further details on antonym modelling, see Paradis, 2009; Paradis & Willners, submitted). However, since entrenchment of form-meaning couplings also plays an important role in the trade-off between memory and reasoning in usage-based modelling of antonymy, we are interested in learning more about the meanings which conventionalize as antonym pairings. The theoretical implication of our approach is that conceptual opposition is the *cause* of lexical relation rather than the other way round, that is, that the opposition is the effect of the lexical relation as the categorical approach would argue. We predict a core of antonymic meanings whose conceptual pre-meaning structure is well-suited for binary opposition and whose lexical correspondences are frequently co-occurring in language use (Jones, 2002, 2007; Murphy, Paradis, Willners, & Jones, 2009; Paradis & Willners, 2007; Willners & Paradis, 2009).

Antonymy and canonicity

Antonyms are at the same time minimally and maximally different from one another. They are associated with the same conceptual domain, but they denote opposite poles/parts of that domain (Croft & Cruse, 2004, pp. 164–192; Cruse, 1986; M. L. Murphy 2003, pp. 43–45; Paradis, 1997, 2001; Willners, 2001). The majority of good opposites, according to speakers' judgements, are adjectives in languages like English, that is, languages which have adjectives. These are also part of the core vocabulary for learners. For instance, the majority of antonyms provided in a learner's dictionary are adjectives (Paradis & Willners, 2007). Most of the pairings are gradable adjectives, either UNBOUNDED expressing a range on a SCALE such as *good – bad*, or BOUNDED expressing a definite 'either-or' mode being able to express totality and partiality such as *dead – alive* (Paradis, 2001, 2008; Paradis & Willners, 2006, 2009), but there are also non-gradable antonymous adjectives such as *male – female*.

Antonymy formed an important part of structuralist models to meaning (Cruse, 1986; Lyons, 1977), in which relations such as antonymy are primitives and meanings of words are the relations they form with other words in the lexical network. Interest in lexical relations faded when the structuralist framework was superseded by conceptual approaches to meaning and the orientation of research interest moved into other areas of semantics, such as event structure and the study of metaphor and metonymy. With the growing theoretical sophistication of Cognitive Semantics and the development of new computational resources, we now see a revival of interest in relations in language, thought and memory. The foundation of relations such as antonymy is still an issue, however. There is no consensus in the literature on the issue of whether antonyms form a set of stored lexical associations, as the structuralists and the Princeton WordNet model propose, or whether the category of antonymy is a context-sensitive, conceptually grounded category of which the members form a prototype structure of 'goodness of antonymy' as conceptual models of meaning would argue (G. L. Murphy, 2002). This section introduces the two contrasting models in that order and then we position ourselves in relation to the types of research that have been used to support their standpoints.

Firstly, the lexical, categorical view of antonymy as proposed by the Princeton WordNet model is shown in Figure 1 (Gross & Miller, 1990, p. 268).

Figure 1 shows the distinction between direct and indirect antonyms, *dry – wet* in this case. The direct antonyms are lexically related, while the indirect ones are linked to the direct antonyms by virtue of being members of their conceptual synonym sets. The direct antonyms are central to the structure of the adjectival vocabulary. Since lexical structure of the Princeton WordNet presupposes the

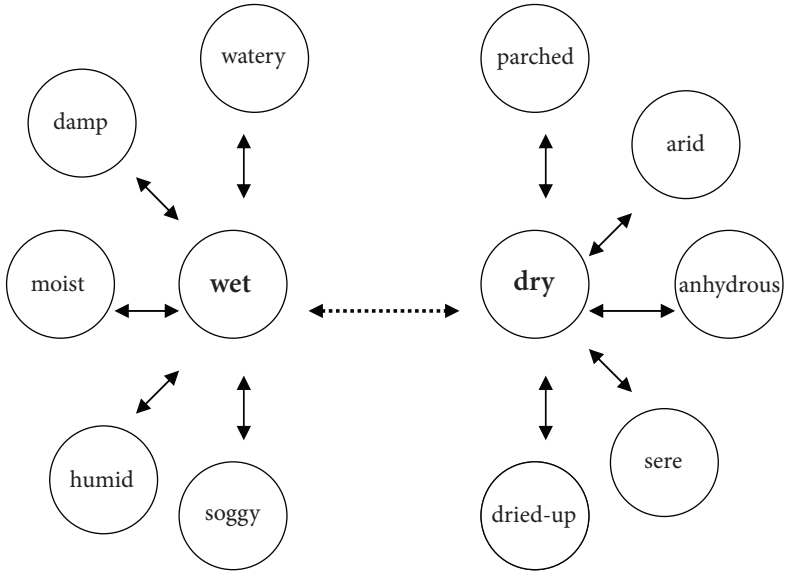

Figure 1. The direct relation of antonymy as illustrated by *wet* and *dry*. The synonym sets of *wet* (i.e., *watery*, *damp*, *moist*, *humid*, *soggy*) and *dry* (i.e., *parched*, *arid*, *anhydrous*, *sere*, *dried-up*) appear as crescents round *wet* and *dry* respectively. They are all indirect antonyms of the direct ones (the figure is adapted from Gross and Miller 1990, p. 268).

existence of direct antonyms, there is a need to make up place-holders for missing members. For instance, *angry* has no partner and therefore UNANGRY is supplied as a dummy antonym. Psycholinguistic indicators that have been used in the literature in support of lexical associations between antonyms include the tendency for antonyms to elicit one another in psychological tests such as free word association (Charles & Miller, 1989; Deese, 1965; Palermo & Jenkins, 1964) and to identify them as opposites at a faster speed (Charles, Reed, & Derryberry, 1994; Gross, Fischer, & Miller, 1989; Herrmann, Chaffin, Conti, Peters, & Fobbins, 1979). For instance, Charles et al. (1994) found that non-canonical antonym reaction times were affected by the semantic divergence between the members of the pair, while reaction times for canonical antonyms were not. Moreover, in semantic priming tests, canonical antonyms have been found to prime each other more strongly than non-canonical opposites (Becker, 1980).

There is, however, evidence that this is an over-simplified means to classify antonyms. Herrmann, Chaffin, Daniel, and Wool (1986) argue that canonicity is a scalar rather than absolute phenomenon. In one of their experiments, Herrmann et al. (1986) asked informants to rate word pairs on a scale from one to five. From the results of their experiment it emerges that there is a scale of 'goodness of antonyms' with scores ranging from 5.00 (*maximize – minimize*) to 1.14 (*courageous*

– *diseased*, *clever* – *accepting*, *daring* – *sick*). Herrmann et al. (1986, pp. 134–135) define antonymy in terms of four relational elements. The first element concerns the clarity of the dimensions on which the pairs of antonyms are based. Their assumption is that the clearer the relation the better the antonym pairing. For instance, according to them the dimension on which *good* – *bad* is based is clearer than the dimension on which *holy* – *bad* relies. The clarity stems from the single component *goodness* for the first pair as compared to the latter pair which they claim relies on at least two pairs, *goodness* and *moral correctness*. In other words, the clearer the dimension is the stronger the antonymic relation. Secondly, the dimension has to be predominantly denotative rather than predominantly connotative. The third element is concerned with the position of the word meaning on the dimensions. In order to be good antonyms the word pairs should occupy the opposite sides of the midpoint, for example, *hot* – *cold*, rather than the same side, for example, *cool* – *cold* (Ogden, 1932; Osgood, Suci, George, & Tannenbaum, 1957). Finally the distances from the midpoint should be of equal magnitude. Each of these elements is a necessary but not a sufficient condition for antonymy, which means that word pairs can fail to conform to the definition of antonymy by failing any one of the four conditions. In the judgement experiment the informants rated the 100 pairs for degree of antonymy on a scale from *not antonyms* (1) to *perfect antonyms* (5). The results show that the degree of antonymy was influenced by the three antonym elements, that is, that the two words are denotatively opposed, that the dimension of denotative opposition is sufficiently clear and that the opposition of two words is symmetric around the centre of the dimension.

Hermmann et al. 1) clarity

2) denotative

3) opposite side of the midpoint

4) equal distance

Similarly, Murphy and Andrew (1993) report on results from a set of experiments on the nature of the lexical relation of antonymy that showed that adjectives are susceptible to conceptual modification. Like Herrmann et al. (1986), they show that opposition is not a clear-cut dichotomy, but a much more complicated and knowledge-intensive phenomenon. In their experiments, antonyms of 14 adjectives from Princeton WordNet were elicited both out of context and in combination with a given noun. They show that the elicited adjectives were not the same across the two conditions, which they take to be evidence of the fact that producing antonyms is a not an automatic association but a knowledge-driven process. The upshot of their study is that antonyms are not lexical relations between word forms, but they have conceptual basis.

Murphy and Andrew (1993) raise four objections against the Princeton WordNet model of antonymy as lexical relations between word forms and not a semantic relation between word meanings. The first objection concerns how antonyms become associated in the first place. One suggestion presented by Charles and Miller (1989) is that they co-occur often. This suggestion is dismissed by Murphy and Andrew on the grounds that it cannot be the final explanation since many other words

co-occur frequently, such as *table* and *chair*, *dentist* and *teeth*. The second objection concerns why they co-occur. If the answer to that is that they co-occur because they are associated in semantic memory, the explanation becomes circular: co-occurrence is caused by the relation and the relation is caused by co-occurrence. Thirdly, if antonymy is just a lexical association, then the semantic component would be superfluous, and this is clearly not the case. On the contrary, the semantic relation is crucial and these semantic properties have to be explained somehow. There are strong theoretical arguments, based on sound empirical evidence, suggesting that word meanings are mentally represented as concepts (G. L. Murphy, 2002, pp. 385–441). In their final discussion, Murphy and Andrew (1993) raise the question of whether there is a place for lexical relations as proposed by Princeton WordNet. Their conclusion is that on the condition that the words happen to be associated, lexical relations may in some cases be pre-stored, but in many other cases they are not. Some lexical relations may be computed from semantic domains where they have never been encountered before, which means that pre-stored lexical links may be an important part of linguistic processing, but they cannot explain the range of lexical relations that can be construed. Murphy and Andrew (1993, p. 318) leave us with this statement and this is where we pick up the baton.

Our study questions both Herrmann et al's (1986) view that antonymy is a completely scalar phenomenon and the categorical view that there is a set of canonical antonyms in language that are represented in the lexicon and another set of non-canonical antonyms that are not represented as pairs in the lexicon, but are understood through a lexicalized pairing as shown in Figure 1. Much like Murphy and Andrew (1993), our hypothesis is that antonymy is conceptual in nature and antonym pairs are always subject to contextual constraints. This is true of all pairings. However, there seems to be a small set of words with special lexico-semantic attraction, and this is where we diverge from Murphy and Andrew. We refer to such pairings as canonical antonyms. They are entrenched in memory and perceived as strongly coupled pairings by speakers. While such strongly conventionalized antonyms form a very limited set, we argue that the majority of adjectives form a continuum from more to less strongly conventionalized pairings across contexts. We also extend the empirical basis for the analysis by including more test items and using both textual and experimental methods. The data, consisting of pairs of words that co-occur in sentences significantly more often than chance would predict, were retrieved from *The British National Corpus* (henceforth the BNC) and used as test items in two different types of experiments: an elicitation experiment and a judgement experiment. In other words, we are drawing on naturally occurring data in text and discourse, antonym production through elicitation and goodness of opposition through speaker judgements of pairings in experimental settings.

The rationale for using a corpus-driven method for data extraction is to make use of natural language production. Previous studies show that textual evidence supports degrees of lexical canonicity. Justeson and Katz (1991, 1992) and Willners (2001) established that members of pairs they perceived to be canonical tend to co-occur at higher than chance rates and that such pairings co-occur significantly more often than other semantically possible pairings (Willners, 2001). Antonym co-occurrence in text is by no means restricted to set phrases such as *the long and the short of it* or *neither here nor there*, but antonym pairs co-occur across a range of different phrases. Indeed, Fellbaum (1995), Jones (2002, 2006, 2007), Mettinger (1994, 1999), Muehleisen and Isono (2009) and Murphy et al. (2009), demonstrate that antonyms frequently co-occur in a wide range of contexts such as *more X than Y*, *difference between X* and *Y*, *X rather than Y*, using both written and spoken corpora.

Treating relations as combinations of conceptual structures, rather than associations between lexical items only, is consistent with a number of facts about the behaviour of relations. Firstly, relations are context dependent and tend to display prototypicality effects in that there are "better" and "less good" instances of relations (Cruse, 1994). In other words, not only is *dry* the most salient and well-established antonym of *wet*, but the relation as such may also be perceived as a better antonym relation than, say, *dry – sweet*, *dry – productive* or *dry – moist*. Also, like categories in general, antonymy is a matter of construals of inclusion, similarity and contrast. The role of antonymy in metonymization and metaphorization is evidence in favour of analogies based on relations of antonymy. At times, new metonymic or metaphorical coinages seem to be triggered by antonym relations. One such example is the coinage of *slow food* as the opposite of *fast food.* Canonicity plays a role in new uses of one of the members of the pair of a salient relation. When a member of a pair of antonyms acquires a new sense, the opposition can be carried into a new domain which is an indication that we perceive the words as related also in that domain. Lehrer (2002) notes that if two lexical items are in a strong relation with one another, the relations can be transposed by analogy to other senses of those words. She illustrates this with *He traded in his hot car for a cold one*. Along the temperature dimension, *hot* contrasts with *cold*, and the relation is carried over to a dimension related to whether the car was legally or illegally acquired. For speakers to be able to understand *cold* in this sense when it is first encountered, they must first of all know the meaning of *hot car* and they must also be familiar with the canonicity of the antonym relation underlying *hot* and *cold* in the temperature dimension. M. L. Murphy (2006) gives examples of the same phenomenon using *black* and *white*. For instance, *black* was in regular use before *white* in expressions such as *black coffee – white coffee*, *black market – white market*, *black people – white people* and *black box testing – white box testing*.

In sum, canonical antonyms are strongly entrenched in memory and language, while the vast majority of potential antonyms are opposites by virtue of their semantic incompatibility when they are used in binary contrast in order to be opposites and weakly associated as lexico-semantic pairings. In spite of the fact that these notions have repercussions for linguistic theories, they have not been defined in a principled way. When researchers distinguish between canonical and non-canonical antonyms for psycholinguistic experiments or when lexicographers decide which relations to represent in their dictionaries or databases (e.g., Princeton WordNet), they do so intuitively and often with unbalanced and irregular results (M. L. Murphy, 2003; Paradis & Willners, 2007; Sampson, 2000).

Aim and hypotheses

The general aim of this study is to gain new insights into the nature of antonymy as a lexico-semantic relation of binary contrast. Our hypothesis is that semantically opposed pairs of adjectives are distributed on a scale from *canonical antonyms* to pairings that are *hardly antonyms at all*. Characteristic of canonical antonyms is that they are conventionalized expressions of the opposing poles. Such pairings are relatively few and they differ significantly from other pairings that are potentially opposable. The great majority of antonym pairings are more loosely connected to one another. Like other categories, the category of antonymy shows prototypicality effects and has internal structure.

Our secondary aim is to make use of a combination of techniques, both textual and psycholinguistic. The principle and method of selection of test items is based on corpus-driven statistical methods described in the next section and the items are subsequently tested in two different types of experiments: a judgement experiment and an elicitation experiment. Our hypothesis is that, irrespective of the technique used, the results will select the same pairings as the best examples of antonyms.

Method of data extraction

As reported above, antonyms co-occur in sentences significantly more often than chance would allow, and some antonym pairs co-occur more often than others. The rationale for the selection of test items for the experiments profits from the statistical findings of co-occurrence of word pairs in textual studies previously carried out (Justeson & Katz, 1991; Willners, 2001). The hypothesis underlying these corpus-driven analyses is that all the words in a corpus are randomly distributed.

Both the above studies prove the null hypothesis wrong, that is, there are word-pairs that occur in the same sentence much more often than expected. Willners (2001) further showed that, antonyms co-occurred significantly more often than all other possible pairings (e.g., synonyms). Using the insights of previous work on antonym co-occurrence as our point of departure, we developed a methodology for selecting data for our experiments. Through the corpus-driven methodology, we could use the corpus to suggest possible candidates for the test set. On the basis of that, we agreed on a set of seven dimensions that we perceived as central meaning dimensions in human communication. We then identified the pairs of antonyms that we thought were the best "opposites" within these dimensions (see Table 1), checking that the antonyms were all represented as direct antonyms in Princeton WordNet.[3] For reasons of methodological clarity, we call this group of antonyms *canonical antonyms* in order to distinguish them from the rest of the antonymic pairings.

The word pairs in Table 1 were then searched in the BNC using a computer program called *Coco* developed by Willners (2001, p. 83) and Willners and Holtsberg (2001). *Coco* calculates expected and observed sentential co-occurrences of words in a given set and their levels of probability. Unlike the program used by Justeson and Katz (1991), *Coco* has the advantage of taking sentence length variations into

Table 1. Seven Dimensions and their Corresponding Canonical Antonym Pairs in English

Dimension	Canonical antonyms
SPEED	slow–fast
LUMINOSITY	light–dark
STRENGTH	weak–strong
SIZE	small–large
WIDTH	narrow–wide
MERIT	bad–good
THICKNESS	thin–thick

Table 2. Sentential Co-occurrences of the Canonical Antonyms in the Test Set

$Word_X$	$Word_Y$	N_X	N_Y	Co	Expct Co	*P*-value
slow	fast	5760	6707	163	9.6609	0.0
dark	light	12907	12396	402	40.0103	0.0
strong	weak	19550	4522	455	22.1076	0.0
large	small	47184	51865	3642	611.9756	0.0
narrow	wide	5338	16812	191	22.4421	0.0
bad	good	26204	124542	1957	816.1094	0.0
thick	thin	5119	5536	130	7.0867	0.0

account. It was confirmed that the seven adjective pairs co-occurred significantly in sentences.

The results of using *Coco* on the seven word pairs in the BNC are shown in Table 2. N_X and N_Y are the numbers of times that the two words occur in the corpus. *Co* is the number of times they co-occur in the same sentence, while *ExpctCo* is the number of times they are expected to co-occur in a way that chance would predict.[4] The figures in the rightmost column show the probability of finding the number of co-occurrences actually observed, or more. The calculations were made under the assumption that all words are randomly distributed in the corpus and the *p-values* are all lower than 0.0001.

Next, all of the synonyms of all 14 adjectives were collected from Princeton WordNet. This resulted in a list of words potentially related to each dimension. For instance, in the SPEED dimension, the list of words contains *fast* and all its synonyms given in Princeton WordNet (n = 64) and *slow* and all its synonyms (n = 39). All of the words in those lists, regardless of their semantic relation, were searched for sentential co-occurrence in the BNC in all possible pairings and orderings on their dimension. The total number of permutations was 68,364. It was established that the seven adjective pairs co-occurred significantly at sentence level as did the pairings of many of their synonyms. Table 3 shows the pairs in the BNC related to the SPEED dimension that co-occur five times or more in the corpus with a *p-value* of 0.0001 or lower.

It is worth noting that the matching of all synonyms within a certain dimension throws up antonym co-occurrences, synonym co-occurrences as well as co-occurrences that might neither be antonyms nor synonyms in any context. For the dimension of SPEED, Table 3 shows that there are significantly co-occurring antonyms, such as *fast – slow*, *rapid – slow*, *quick – slow* and significantly co-occurring synonyms, such as *fast – quick*, *fast – rapid*, *boring – dull* and *sudden – swift* as well as pairs that might neither be antonyms, nor synonyms in any context but nevertheless co-occur significantly, such as *dense – hot*. Appendix A shows the top ten pairings for each of the seven dimensions across all the 68,364 possible permutations. All seven canonical pairs (originally chosen to represent the dimension) are found in this very limited list. Four of them appear at the very top of their dimensional field, namely *fast – slow, strong – weak, small – large* and *bad – good* when sorted according to falling number of sentential co-occurrences. *Dark – light* is on even footing with, for example, *black – white*, *blue – white* and *black – dark*, while the sentential co-occurrence of *narrow – wide* is lower than word pairs that rather belong to the more complex dimension of SIZE, such as *big – large*.[5]

Table 3. Sentential Co-occurrences of Ssynonyms of Fast and Slow in the BNC with p-Value $\leq 10^{-4}$, Co-occurring more than 5 Times

$Word_X$	$Word_Y$	$N1_X$	$N2_Y$	Co	Expct Co	P-value
fast	slow	6707	5760	163	9.6609	0.0
rapid	slow	3526	5760	54	5.0789	0.0
quick	slow	6670	5760	39	9.6076	0.0
fast	quick	6707	6670	34	11.1871	0.0
firm	smooth	6157	3052	34	4.6991	0.0
fast	rapid	6707	3526	29	5.9139	0.0
gradual	sudden	1066	3920	22	1.0450	0.0
gradual	slow	1066	5760	22	1.5355	0.0
gradual	immediate	1066	6104	18	1.6272	0.0
boring	dull	1669	1837	17	0.7667	0.0
dense	hot	1060	9445	15	2.5036	0.0
sudden	swift	3920	920	14	0.9019	0.0
dull	slow	1837	5760	14	2.6460	0.0
instant	quick	1638	6670	13	2.7322	0.0
lazy	slow	819	5760	10	1.1797	0.0
lazy	stupid	819	3234	9	0.6624	0.0
slow	tedious	5760	543	9	0.7821	0.0
delayed	immediate	450	6104	9	0.6869	0.0
fast	high-speed	6707	359	8	0.6021	0.0
slow	sluggish	5760	220	8	0.3169	0.0
smooth	swift	3052	920	7	0.7022	0.0
faithful	loyal	1005	1320	7	0.3317	0.0
dense	smooth	1060	3052	7	0.8090	0.0
dumb	stupid	755	3234	7	0.6106	0.0
boring	tedious	1669	543	6	0.2266	0.0
fast	speeding	6707	104	6	0.1744	0.0

The test items

From our long-list of co-occurring pairs, the next step was to derive a test set of pairs for use for the experiments. As stated in the aim, the hypothesis we are testing is that oppositeness is a continuum, and opposite pairings are distributed on a scale from well-established canonical antonyms to pairings that are not well-established as antonyms. We also predict that there is a group of strongly

conventionalized antonym pairings that differ significantly from the rest of the antonyms and that the potentially opposable synonyms and unrelated pairings in turn differ significantly from all antonyms. It was important to include synonyms since, like antonyms, the relation of synonymy also relies on both similarity and difference. We included in the test set the antonyms listed in Table 1 that were all found to be well-established, intuitively as well as in terms of sentential co-occurrence. We call them canonical antonyms to distinguish the two antonym conditions. Using Princeton WordNet and dictionaries we tagged the significantly co-occurring word-pairs according to semantic relation: antonym, synonym or unrelated. An additional criterion to qualify as antonymous in the test set was that they should all be compatible with scalar degree modifiers such as *very*. The reason for the delimitation to scalar antonyms was that we wanted the test set to be as homogenous as possible. For each dimension, we selected two pairs of antonyms, two pairs of synonyms for each dimension and one pair of co-occurring adjectives that did not appear to be related at all, but which still co-occurred significantly with a *p-value* at 0.0001. Table 4 shows the complete set of pairs retrieved by this method from the BNC: 42 pairs in total.

In addition to the co-occurring pairs, our test set includes a subset of the 100 pairs from Herrmann et al.'s (1986) data set. Their experiment includes mainly adjectives but also verbs and nouns and it shows that there is a scale of 'goodness of antonyms'. Since our study focuses on scalar adjectives, we have excluded the

Table 4. The Cco-occurring Test Items Retrieved from the BNC (for a Definition of Synonymy see Cruse, 1986, pp. 265–290)

Canonical antonyms	**Antonyms (2 per canonical pair)**	**Synonyms (2 per canonical pair)**	**Unrelated**
slow–fast	slow–sudden gradual–immediate	slow–dull fast–rapid	hot–smooth
light–dark	gloomy–bright pure–black	light–pale dark–grim	clean–easy
weak–strong	delicate–robust tender–tough	weak–feeble strong–firm	slight–soft
small–large	modest–great small–enormous	small–tiny large–huge	heroic–young
narrow–wide	narrow–open limited–extensive	narrow–slender wide–broad	bare–slender
bad–good	bad–mediocre evil–good	bad–poor good–healthy	big–white
thin–thick	lean–fat rare–abundant	thin–fine thick–heavy	pale–slim

Table 5. The Sample of Eleven Antonym pPairs of Decreasing Degrees of Goodness of Antonymy from Herrmann et al's (1986) Test Items

Pairs	Score	Category
beautiful–ugly	4.90	Canonical antonyms
immaculate–filthy	4.62	Antonyms
tired–alert	4.14	Antonyms
disturbed–calm	3.95	Antonyms
hard–yielding	3.28	Antonyms
glad–irritated	3.00	Antonyms
sober–exciting	2.67	Antonyms
nervous–idle	2.24	Unrelated
delightful–confused	1.90	Unrelated
bold–civil	1.57	Unrelated
daring–sick	1.14	Unrelated

verbs, nouns and non-scalar adjectives from Herrmann et al.'s list, leaving us with 63 scalar adjectives that had not already qualified for our test set. We sorted the word pairs according to decreasing scores and then picked out every sixth pair counting from the bottom of the list. This method left us with the eleven word pairs shown in Table 5.

The left column of Table 5 shows the 11 pairs selected from Herrmann et al.'s (1986) data set. The column in the middle shows the mean scores given by the experiment participants, with 5 being the highest possible level of opposition, and the right column indicates our categorization of the pairs. The word pairs were categorized according to the same principles as those in Table 4. All the pairs with a score lower than 2.50 we relegated to the group of unrelated pairings. The entire test set is presented in Table 6 below.

Judgement experiment

This section describes the judgement experiment in which participants were asked to evaluate word pairings in terms of how 'good' they thought each pair is as a pair of opposites. The experiment was carried out through a computer interface. The design of the screen is shown in Figure 2.

The participants were presented with questions of the form: *How good is* X – Y *as a pair of opposites*, as shown in Figure 2. The question was formulated using *good* (not *bad*) in order for the participants to understand the questions as impartial

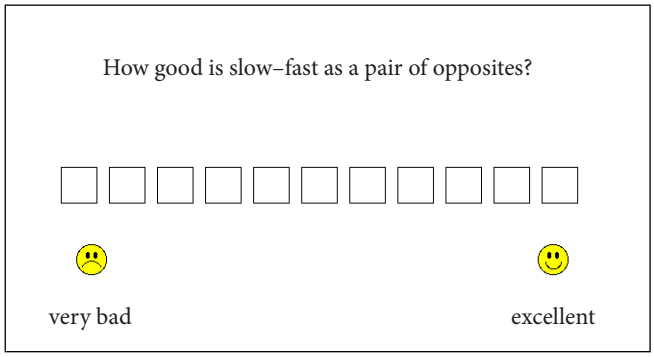

Figure 2. Screen snapshot: An example of a judgement task in the online experiment.

how-questions. *How bad is fat – lean as a pair of opposites* presupposes 'badness'. This principle is consistent with Lehrer's (1985, p. 400) markedness properties for members of antonym pairs. For instance, *good* in *How* GOOD *is it?* with the principal tone on *good* carries no supposition as to which part of the scale of MERIT is involved (Cruse, 1986). The end-points of the scale were designated with both icons and text. On the left-hand side there is 'a frowning face' and underneath the frowning face it says *very bad*. On the right-hand side is 'a smiling face' and the text underneath is *excellent*. The task of the participants was to tick a box on a scale consisting of eleven boxes. Our predictions, mostly underpinned by the theoretical statements made in the introduction, were as follows.

- The eight test pairings categorized as canonical antonyms will receive an average score that is significantly higher than the other word pairs in the test set.
- The sequence of the antonyms, that is, *Word 1 – Word 2* versus *Word 2 – Word 1*, will not significantly affect judgements of 'goodness' in canonical antonyms and antonyms.
- There will be significant differences between the judgements about canonical antonyms, antonyms, synonyms and unrelated pairings, with canonical antonyms at one extreme, unrelated at the other extreme, and antonyms and synonyms in between.
- The response times for the judgements about goodness of oppositeness will be significantly faster for canonical antonyms than for antonyms. The response times for the judgements for antonyms will be significantly faster than for synonyms and unrelated.

Stimuli. The stimuli in the judgement experiment were presented in pairs. The test items were automatically randomized for each participant. The sequence of the individual pairs was designed so that half of the participants were given the test items in the order *Word 1 – Word 2,* while the other half were presented the

Table 6. The Complete Set of Test Items for the Judgement Experiment

Canonical antonyms		Antonyms		Synonyms		Unrelated	
Word 1	**Word 2**	**Word 1**	**Word 2**	**Word 1**	**Word 2**	**Word 1**	**Word 2**
slow	fast	slow	sudden	slow	dull	hot	smooth
light	dark	gradual	immediate	fast	rapid	clean	easy
weak	strong	gloomy	bright	light	pale	slight	soft
small	large	pure	black	dark	grim	heroic	young
narrow	wide	delicate	robust	weak	feeble	bare	slender
bad	good	tender	tough	strong	firm	big	white
thin	thick	modest	great	small	tiny	pale	slim
ugly	beautiful	small	enormous	large	huge	nervous	idle
		narrow	open	narrow	slender	delightful	confused
		limited	extensive	wide	broad	bold	civil
		mediocre	bad	bad	poor	daring	sick
		evil	good	good	healthy		
		lean	fat	thin	fine		
		rare	abundant	thick	heavy		
		filthy	immaculate				
		calm	disturbed				
		tired	alert				
		hard	yielding				
		sober	exciting				
		irritated	glad				

words in reverse order, that is, *Word 2 – Word 1*, as listed in Table 6. We call the two conditions non-reversed (*Word 1 – Word 2*) and reversed (*Word 2 – Word 1*) respectively.

The principle underlying the ordering of the pairs is one of polarity. The meaning of *Word 1* denotes 'little' of the property expressed by both members of the pairs, and inversely *Word 2* denotes 'much' of the property, for example, *slow* (little) – *fast* (much) of the property SPEED. In the cases where there is no clear pattern of "lacking" versus "having" of the property denoted, the meanings are associated with negative and positive evaluation, for example, *bad – good*. The principle we used for them was that the word meanings associated with negative evaluation was aligned with "little of the property", that is, *Word 1*, and the positively oriented word meanings were aligned with 'much of the property', for example, *bad* (negative) – *good* (positive) of the property of MERIT. These distinctions apply to the two sets of antonyms only, that is, canonical antonyms and antonyms and are crucial for the verification of our prediction that the sequence of the antonyms is not of any importance. The ordering predictions are not applicable to the synonym and unrelated categories.

Participants. Fifty native speakers of English participated in the judgement test, none of whom would also participate in the elicitation test. The informants were

students, faculty, administrative staff, caretakers, bus drivers and other visitors to Sussex University (where the experiment was conducted). Thirty-two of the participants were women and 18 were men. All had English as their first language. Six participants had a parent with a native language other than English (French, Polish, Hebrew, Welsh and Greek) and one participant's parents both spoke Luganda.

Procedure. The judgement experiment was performed using E-prime as experimental software.[6] The participants were presented with a new screen for each word pair (see Figure 2). The task of the participants was to tick a box on a scale consisting of eleven boxes. The screen immediately disappeared upon clicking, which prevented the participants from going back and changing their responses. Between each judgement task a screen with only an asterisk was presented. Each participant completed two test trials before the actual judgement test of the 53 test items. The purpose of the study was revealed to the participants in the instructions.

Participant ratings and response times were subjected to one-way ANOVAs (F_1 and F_2 analyses), followed by pairwise comparisons using Bonferroni corrections. In cases where parametrical tests were potentially problematic, that is, when assumptions of homogeneity or sphericity were violated, we also performed a corresponding nonparametric test. The results of the nonparametric tests were the same as those of the parametric ones and are therefore not reported below.

As has already been mentioned, the judgement experiment was divided into two parts: 25 participants were given the test set as non-reverse (*Word 1 – Word 2*, e.g., *slow – fast*) and 25 participants were given the test set in the reverse order (*Word 2 – Word 1*, e.g., *fast – slow*). This was done to control for whether sequence influenced the results in any way. In both these experiments, the non-reverse and the reverse, a subject analysis and an item analysis were performed. The factors involved were sequential ordering, category (canonical antonyms, antonyms, synonyms and unrelated) and the interaction between sequential ordering and category.

Results of judgement experiment

This section reports on the results of the judgement experiment. Three different aspects were measured: (i) whether the differences between the four categories (canonical antonyms, antonyms, synonyms and unrelated pairings) were significant and their levels of strength of opposability, (ii) ordering of presentation of the words on the screen, and (iii) response time. Before going into details about the results, it should be mentioned that there were 10 zeroes among the responses due to the fact that some participants had ticked outside the boxes on the scale. They were not excluded in the analysis, since they are too few to affect the results in any way.[7]

Canonical antonyms, antonyms, synonyms and unrelated pairings. Table 7 shows that the mean response values for the canonical antonyms were 10.63, the antonyms 7.66, the synonyms 1.63 and the unrelated 1.77, see Table 7. The standard deviation is much larger for the antonyms than for the other three categories; that is, the informants agree less strongly about the judgement of the antonyms than of the other categories of word pairs. It is 3.23 for the antonyms, while it varies between 1.20 and 1.51 for the other categories. Since the order did not have any effect, see the section on *Sequential ordering* below, we collapsed all data here.

Table 7. Mean Responses for Canonical Antonyms, Antonyms, Synonyms and Unrelated Word Pairs

Category	Mean	Std. Deviation
Canonical antonyms	10.63	1.20
Antonyms	7.66	3.23
Synonyms	1.63	1.51
Unrelated	1.77	1.30

The results for each of the word pairs in the judgement test are presented in Table 8, sorted according to falling response means. The results show that the participants were in agreement about the canonical antonyms. The top eight word pairs of antonyms, which we call canonical antonyms, yield mean responses over the subjects between 10.30 and 10.82. The next 19 word pairs were classified as antonyms before the test and their mean responses vary between 3.00 and 10.24. The variation is also reflected in the standard deviations across all the individual antonym pairs. The standard deviations are much larger for the antonyms than for the word pairs in the other three categories.

One of our main hypotheses was that the variation in strength of lexico-semantic couplings would yield significant differences between four different groups of word pairs: canonical antonyms, antonyms, synonyms and unrelated word pairs. We also expected very strong agreement among the participants concerning the canonical antonyms. We did find significant differences between the judgements of the canonical antonyms, the antonyms and the other two categories, synonyms and unrelated word pairs. This is shown in Table 8 where the top eight pairs are the antonyms that we chose to call canonical, followed by the other 19 pairs of antonyms. Synonyms and unrelated word pairs are mixed in the bottom part of the table. Contrary to what we predicted, there was no significant difference between synonyms and unrelated word pairs.

According to the repeated-measures ANOVA for the subject analysis and the item analysis, the differences between the canonical antonyms and antonyms as well as

Table 8. Mean Responses, Standard Deviations and Categorization of the Word Pairs in the Test Set

Word 1	Word 2	Mean response	Std. Deviation	Category
weak	strong	10.82	0.44	C
small	large	10.82	0.39	C
light	dark	10.68	1.35	C
narrow	wide	10.66	0.72	C
thin	thick	10.66	0.77	C
bad	good	10.64	1.21	C
slow	fast	10.42	1.73	C
ugly	beautiful	10.30	1.93	C
evil	good	10.24	1.65	A
limited	extensive	9.92	1.18	A
delicate	robust	9.84	1.11	A
lean	fat	9.76	1.76	A
rare	abundant	9.66	1.71	A
small	enormous	9.30	2.01	A
filthy	immaculate	9.30	2.17	A
calm	disturbed	9.30	1.69	A
tired	alert	9.10	1.22	A
tender	tough	9.08	2.11	A
gradual	immediate	8.78	1.84	A
hard	yielding	7.60	2.65	A
slow	sudden	6.44	3.04	A
narrow	open	5.82	3.32	A
sober	exciting	5.38	2.64	A
irritated	glad	5.20	2.84	A
modest	great	4.72	2.98	A
pure	black	3.04	2.45	A
mediocre	bad	3.00	1.98	A
bold	civil	2.88	1.96	U
idle	nervous	2.40	1.80	U
delightful	confused	2.28	1.59	U
bad	poor	2.26	2.42	S
good	healthy	1.96	1.76	S
wide	broad	1.88	2.44	S
light	pale	1.76	1.62	S
small	tiny	1.66	1.80	S
narrow	slender	1.64	1.72	S

Table 8. *(continued)*

Word 1	Word 2	Mean response	Std. Deviation	Category
slight	soft	1.58	0.84	U
slow	dull	1.58	1.07	S
hot	smooth	1.58	1.30	U
thin	fine	1.58	1.14	S
bare	slender	1.56	1.09	U
heroic	young	1.54	0.97	U
daring	sick	1.52	0.89	U
strong	firm	1.48	0.79	S
fast	rapid	1.48	1.61	S
dark	grim	1.46	0.71	S
clean	easy	1.46	0.65	U
thick	heavy	1.44	0.84	S
pale	slim	1.40	0.73	U
weak	feeble	1.38	0.64	S
large	huge	1.32	0.84	S
big	white	1.32	0.65	U

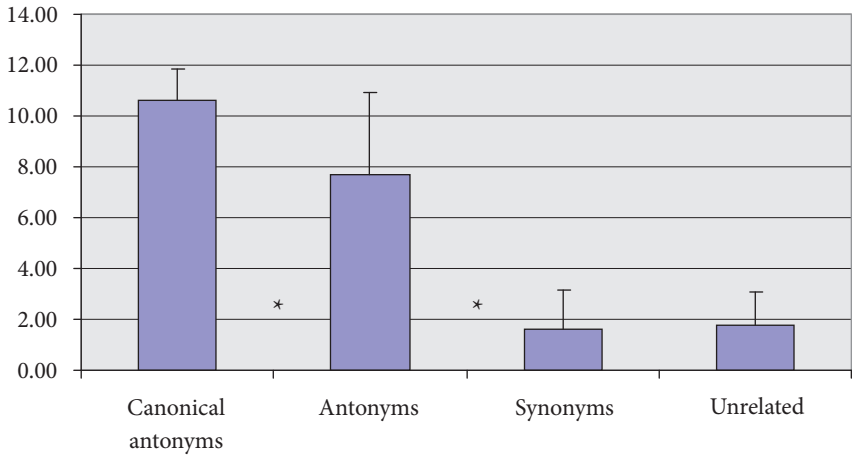

Figure 3. Mean responses for canonical antonyms, antonyms, synonyms and unrelated word pairs.

between antonyms and the two other categories (synonyms and unrelated) were significant both in the subject analysis, $F_1(3, 147) = 1625.775$, $p < 0.001$, and in the item analysis, $F_2(3,48) = 95.736$, $p < 0.001$. The post-hoc comparisons suggested that the four conditions form three subgroups: (1) canonical antonyms, (2) antonyms and (3) synonyms and unrelated.

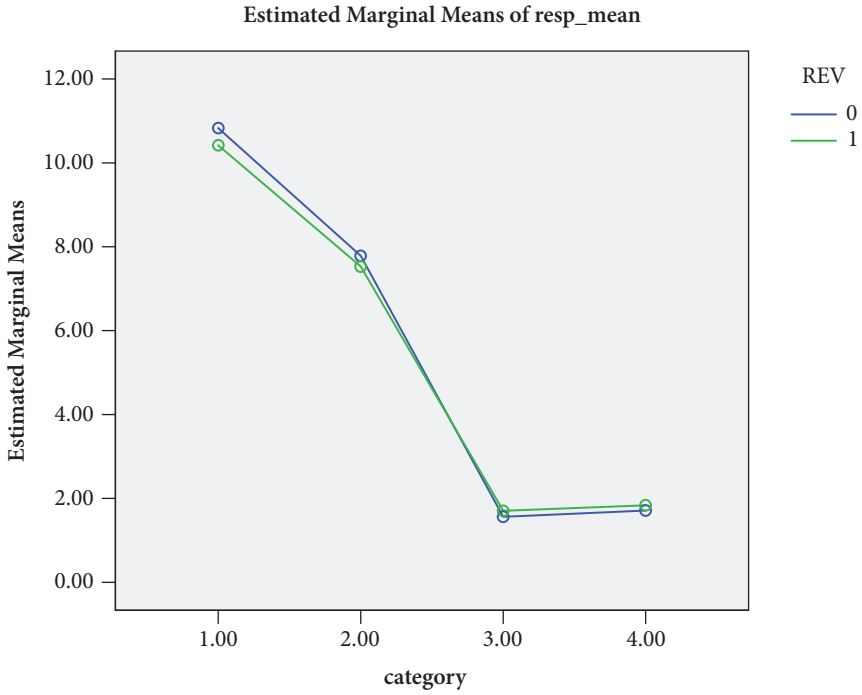

Figure 4. Sequential ordering: There is no significant difference between the mean answers of the two test batches.

Sequential ordering. The statistical analysis corroborates our prediction that sequential ordering does not have any effect on the judgements, yielding very low F-values: $F_1(1, 48) = 0.558$, $p = 0.459$; $F_2(1, 96) = 0.101$, $p = 0.751$. The interaction between the sequential ordering and category shows no effect either: $F_1(3, 144) = 1{,}582$, $p = 0.196$; $F_2(3, 96) = 0.186$, $p = 0.906$. Category on the other hand has an effect: $F_1(3, 144) = 1645{,}082$, $p < 0.001$; $F_2(3, 96) = 187{,}449$, $p < 0.001$. Figure 4 shows that the two test batches (marked with REV = 0 and REV = 1) follow the same pattern. Since the direction does not have an impact on the results, the data for the two directions are treated as one batch as in 5.1.1 and in 5.1.3.

Response times. Because of the fact that the experiment was self-paced, that is, the participants could use the time they needed for each judgement, we cannot draw any far-reaching conclusions about the time it took for the participants to make their decisions. Still, it may be of some interest to note that the response times varied greatly across the conditions (see Table 9 and Figure 5). The overall effect was significant in the subject analysis, $F_1(3, 147) = 27.256$, $p < 0.001$, and in the item analysis, $F_2(3, 48) = 23.733$, $p < 0.001$. According to the post hoc test, the canonical antonyms take significantly shorter to process than unrelated word pairs and the

Table 9. Mean Response Times for Canonical Antonyms, Antonyms, Synonyms and Unrelated Word Pairs

Category	Mean response time	Std. Deviation
Canonical antonyms	4303.5933	1699,90
Antonyms	7648.5294	3517,32
Synonyms	5446.1427	2504,50
Unrelated	6381.9891	3080,64

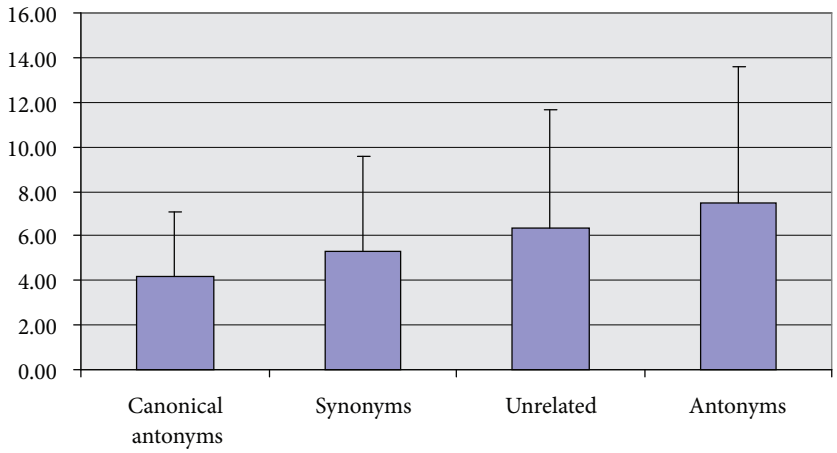

Figure 5. Mean response times for canonical antonyms, antonyms, synonyms and unrelated word pairs.

antonyms. The antonyms take significantly longer to process than the synonyms. There is no significant difference between the response times of the synonyms and the unrelated word pairs. The main result of the analysis of the response times is that the canonical antonyms are significantly faster to process than the antonyms.

Elicitation experiment

This section reports on the design and the results of the elicitation experiment. The participants were asked to provide the best opposite for all the individual test items. Our predictions were as follows.

- The 16 test items from the set of canonical antonyms will elicit only one another.
- The other test items will elicit varying numbers of antonyms — the better the antonym pairing, the fewer the number of elicited antonyms.

– The elicitation experiment will produce a curve from high participant agreement (only one antonym suggested) to low participant agreement (many suggested antonyms).

The predictions here rely on the basic theoretical assumption that canonical antonyms are strongly entrenched lexico-semantic couplings and for that reason they strongly prefer one anothers' company.

*Stimuli and procedure.*The test set for the elicitation test involves the same individual adjectives as in the judgement experiment (see Table 6). Some of the individual adjectives occur in more than one pair, that is, they might occur only once, twice or three times. For instance, *small* occurs three times and *large* occurs twice. All doublets and triplets were removed from the elicitation test set, which means that *small* and *large* occur once in the elicitation experiment. In total, the experiment contains 85 randomized seed words. The words were presented to all participants in the same order on the test occasion. The participants were asked to write down the best opposites they could think of for each of the 85 stimulus words in the test set. For instance,

> The opposite of LITTLE is _____
> The opposite of DELIGHTFUL is _____

The experiment was performed using paper and pencil and the participants were instructed to do the test sequentially that is, to start from word one, work their way through the experiment and not go back to check or change anything. We did not control for time, but the participants were asked to write the first opposite word that came to mind (see the instructions to the participants in Appendix B). Each participant also filled in a cover page with information about name, sex, age, occupation, native language and parents' native language(s). All the responses were then coded into a database using the stimulus words as anchor words.

Participants. Fifty native speakers of English participated in the elicitation experiment: 36 women and 14 men. The experiments were carried out at the University of Sussex in England and at Lund University and Växjö University in Sweden. The informants were students, faculty, administrative staff, caretakers and other visitors. All were native speakers of English and none had participated in the judgement experiment.

Results of elicitation experiment

The data were analyzed with respect to the total distribution of responses across participants, omitted responses were identified and strength of bidirectional elicitation measured through cluster analysis.

Distribution of participant responses. The result of the elicitation experiment indicates that a continuum of lexical association exists between antonym pairs. The results in Appendix C have been listed in order of the number of participants' responses — that is, response diversity. At the top of the list we find the test words for which all participants suggested one and the same antonym, given in brackets: *bad* (*good*), *beautiful* (*ugly*), *clean* (*dirty*), *heavy* (*light*), *hot* (*cold*), *poor* (*rich*) and *weak* (*strong*). The test items for which the participants suggested two opposites are then listed, for example, *narrow* (*wide*, *broad*) and *slow* (*fast*, *quick*) and then the stimulus words with three different answers and so on. The very last item is

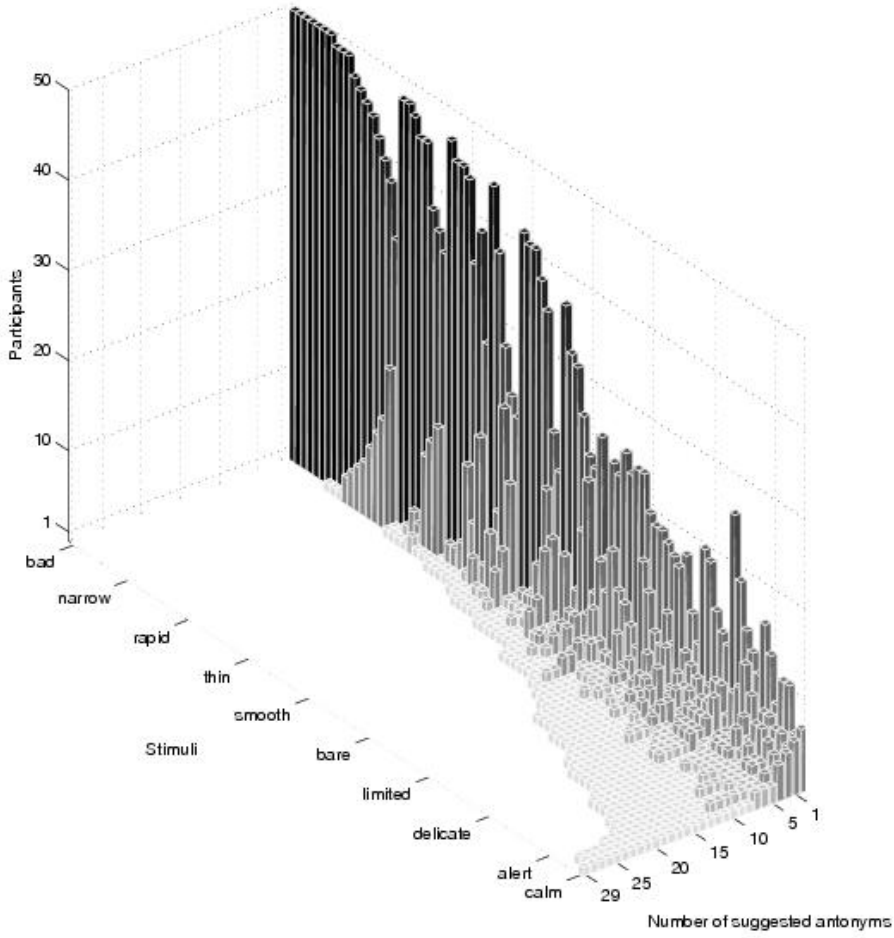

Figure 6. The distribution of English antonyms in the Elicitation experiment. The Y-axis gives the test items, with every tenth test item is written in full, the X-axis gives the number of suggested antonyms across the participants given on the Z-axis.

calm, for which 29 different antonyms were suggested by the 50 participants. The shape of the list of elicited antonyms across test items in Appendix C strongly suggests a scale of canonicity from very good matches to test items with no preferred partners.

While Appendix C gives all the elicited antonyms across the test items, it does not provide information about the scores for the various individual elicited responses. Figure 6 gives the complete three-dimensional picture of the responses. The X-axis gives the total number of the antonyms suggested across each test word. The Y-axis shows all the test items of which every tenth word is supplied along the axis. The Z-axis shows the number of given participant responses participants per antonym. The bars represent the various elicited antonyms in response to the test items. The height of the bars indicates the number of participants who suggested the antonym in question. There is a gradual decrease across stimuli in participant agreement of the best antonym for a given word. The low bars at the front represent a single antonym suggested by one experiment participant. Finally, Figure 7 gives an example of the numbers of the responses to the test item *pale*. *Dark* was suggested by 21 participants and therefore was the "best" antonym followed by *dull*, *dim*, *gloomy*, *stupid* and *obscure* with steadily decreasing numbers.

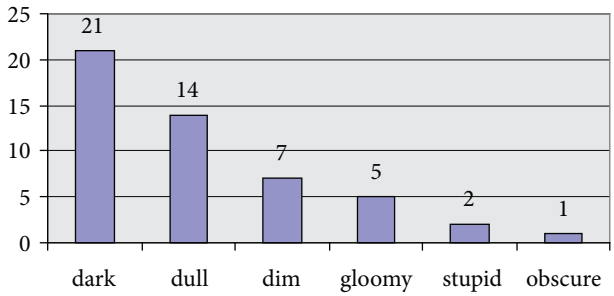

Figure 7. The participants' responses to *pale*.

Omitted responses. Although the participants in the elicitation experiment were asked to give opposites for all words, not all participants responded to all the test items. Out of a total of 4 250 responses (85 test words x 50 participants), participants failed to supply an antonym for 94 test items. The majority of those 94 test items were among the test items that attracted the highest number of suggested antonyms. Table 10 shows the test words for which all participants suggested the best antonym and the test items for which responses were omitted by at least one participant.

The two groups are about the same size: there are 41 words in the group where all participants answered and 44 words in the group with omitted answers. The mean of the number of suggested antonyms in the left column of Table 10, that is,

Table 10. Test Words for which all Participants Suggested an Antonym and Test Words for which some of the Participants did not Provide an Antonym

Test words with 50 responses	Number of suggested antonyms	Test words with omitted responses	Number of suggested antonyms
bad	1	young	1
beautiful	1	white	2
clean	1	light	2
heavy	1	dark	2
hot	1	thick	3
poor	1	sober	5
weak	1	sick	5
black	2	fat	6
fast	2	rare	6
narrow	2	feeble	6
slow	2	broad	6
soft	2	lean	8
good	2	heroic	8
hard	2	glad	8
open	2	bare	8
big	2	gradual	9
easy	2	slim	9
large	3	sudden	10
rapid	3	gloomy	11
small	3	pale	13
ugly	3	nervous	13
exciting	3	limited	13
strong	4	robust	13
wide	4	fine	14
evil	4	abundant	14
thin	4	pure	14
filthy	5	immaculate	14
huge	5	civil	15
enormous	6	extensive	16
dull	6	grim	16
bright	6	slender	17
smooth	6	delicate	17
healthy	7	immediate	17
tiny	7	modest	17
tough	9	firm	18
tired	10	daring	19
idle	11	confused	19
tender	12	bold	19
great	18	mediocre	19
alert	22	yielding	20
calm	29	irritated	21
		disturbed	23
		slight	26
		delightful	27

the test items for which all 50 participants suggested antonyms, is 5.3 and the median is 3. In comparison, the arithmetic mean of the number of suggested antonyms in the column to the right (where the participants failed to suggest antonyms) is 12.5 and the median is 13. This is an indication that the test items for which participants abstained from providing an antonym elicited many more suggestions on average than did the ones that where all participants suggested an antonym.

Bidirectionality. In addition to the distribution of the responses for all the test items across all the participants, we also investigated to what extent the test items elicited one another in both directions. For instance, 50 participants gave *strong* as an antonym of *weak*, *good* for *bad*, *ugly* for *beautiful* and *light* for *heavy*, but the pattern was not the same in the other direction. This is part of the information in Appendix C and Figure 6. For the test items that most speakers of English intuitively deem to be good pairs of antonyms, the strong agreement held true in both directions, not always at the level of a one-to-one match, but a one-to-two or one-to-three (see Appendix C). For example, while 50 participants supplied *good* as the best opposite of *bad*, two antonyms were suggested for *good*: *bad* by 42 participants and *evil* by 8 participants, as shown in Figure 8. This points to the possibility that there is a stronger relationship between *good* and *bad* than between *good* and *evil*. *Dark* and *heavy* were given for *light*, and *beautiful*, *pretty* and *attractive* for *ugly*, which shows that there are differences with respect to which test items of a given pair the participants were confronted with (cf. the results of the judgement experiment).

In summary, Figure 6 shows that the strongly related pairs elicit only one or two antonyms, while there is a steady increase in numbers of preferred antonyms the further we move toward the right-hand side of the figure.

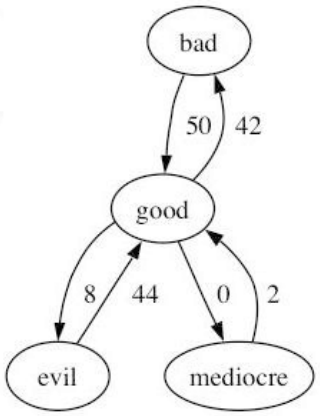

Figure 8. Relations between good, bad, evil and mediocre based on the elicitation experiment. The number of responses is marked by each arrow.

Cluster analysis. In order to throw light on the strength of the conventionalized oppositeness in the elicitation experiment, a cluster analysis of strength of antonymic affinity between the lexical items that elicited one another in both directions was performed. More precisely, a hierarchical agglomerative cluster analysis using the Ward amalgamation strategy was performed on the subset of the data that were bidirectional. Agglomerative cluster analysis is a bottom-up method that takes each entity, that is, in this case each antonym pairing, as a single cluster to start with and then builds larger and larger clusters by grouping together entities on the basis of similarity. It merges the closest clusters in an iterative fashion by satisfying a number of similarity criteria until the whole dataset forms one cluster. The advantage of cluster analysis is that it highlights associations between features as well as the hierarchical relations between these associations. It is not a confirmatory analysis but a useful tool for exploratory purposes (Divjak & Gries, 2008; Glynn, Geeraerts, & Speelman, 2007, Gries & Divjak, in press).

It is important to note that for this experiment only responses that were also test items were eligible as candidates for participation in bidirectional relations. This means that not all of the pairings suggested by the participants were included in the cluster analysis. The participants were free to suggest any word they thought was the best antonym. For instance, *quick* was considered the best antonym of *slow* by five of the participants (as compared to 45 for *fast*), but since *quick* was not included among the stimulus items, the pairing was not included in the cluster analysis. The results of the cluster analysis are, however, comparable to the results of sentential co-occurrence of antonyms in the corpus data and the results of the judgement experiment of the word pairs comprised in all three analysis components.

Figure 9 shows the dendrogram produced by the cluster analysis. It is a hierarchical structure of clusters with two branches at the top. The left branch hosts Cluster 1 and Cluster 2 and the right branch Cluster 3 and Cluster 4. The closeness of the fork to the sub-clusters reveals that there is a closer relation between Clusters 3 and 4 than between Clusters 1 and 2.

The actual pairings are given in the boxes at the end of the branches in Figure 9. There are fewer pairs at the end of the left-most branches than at the ends of the branches on the right-hand side. Six of the word pairs in Cluster 1 were included in the test set as canonical antonyms (subscripted *c* in Figure 9): *bad – good*, *beautiful – ugly*, *light – dark*, *narrow – wide*, *weak – strong* and *fast – slow*. The rest of the word pairs in Cluster 1 were not included as pairs in the experiment. They appear in Cluster 1 because the seed word was in the data set and the pairings were suggested by most (nearly all) of the participants. Cluster 2 includes two word pairs featured in the test set as canonical antonyms: *large – small* and *thick – thin*. Both of these word pairs have different combinatorial preferences (cf. *big –*

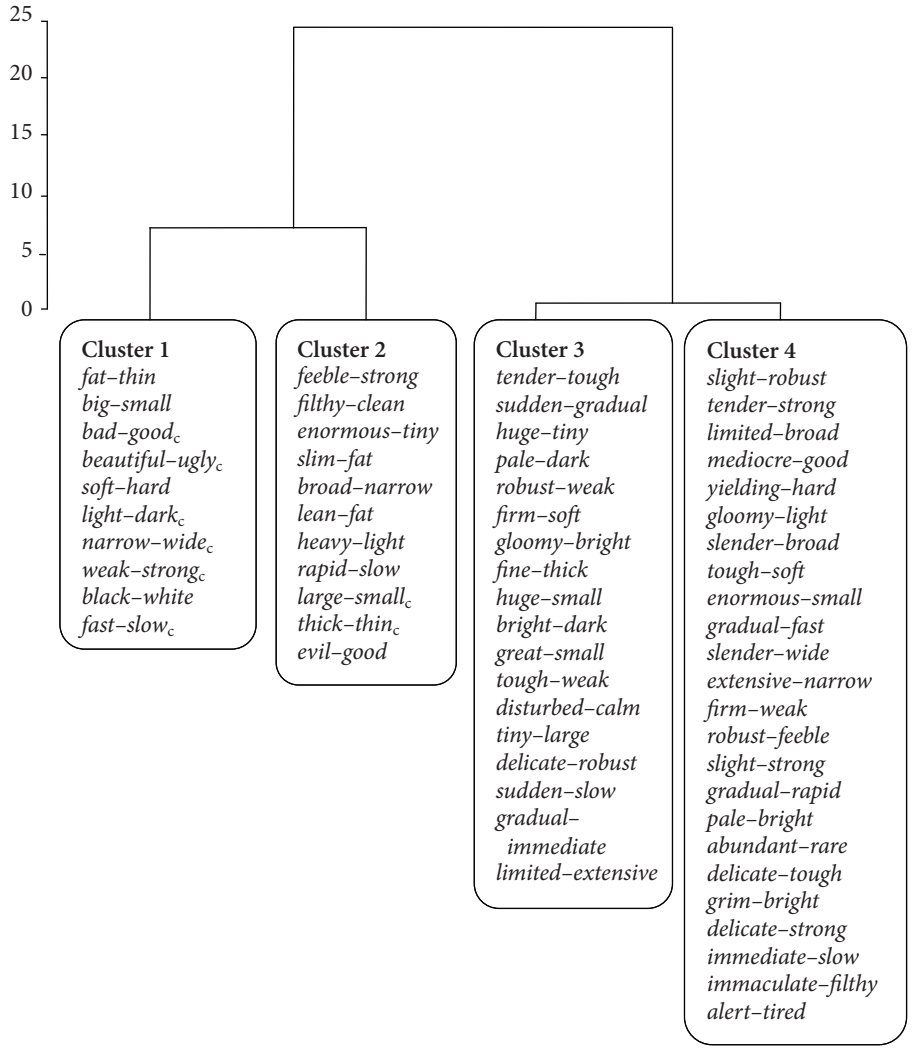

Figure 9. Dendrogram of the bidirectional data.

large – small – little and *thick – thin – fat – fit – slim*) or are clearly polysemous like *heavy – light* (cf. *dark – light*). In spite of the fact that the rest of the word pairs in Cluster 2 seem to be good examples of opposability, they were not among the pairings that we deemed canonical antonyms in the design of the test set, for example, *feeble – strong*, *filthy – clean*, *enormous – tiny*.

Discussion

Our main hypothesis was that opposite pairs of adjectives are distributed on a scale from *canonical antonyms* to *weakly antonymic couplings*. While the number of less strongly associated antonyms is assumed to be, in principle, infinite in that almost any two meanings can make a word pair contrastive given a suitable context, the number of strongly conventionally associated pairings was assumed to be very limited. We expected the category of antonymy to have a number of prototypical antonyms which are associated by semantic opposition as well as by linguistic convention. In contrast to the categorical view, we also predicted that like any other category, the category of antonymy has internal structure, that is, a number of strongly related lexical pairings and a steady increase in partners towards the borders of the category.

We found it important to carry out the retrieval of data for the experiments in a way that involved as little interference from the members of the research group as possible. For that reason we opted for a corpus-driven methodology of data extraction with minimum involvement of intuitive judgements by native speakers and minimum assumptions made by the research team.[8] The advantage of using a corpus-driven method of data extraction was that this technique also yields results in itself. The corpus evidence showed that the patterns for pairs that demonstrate strongly significant co-occurrence in text coincide with the patterns for strong and less strong canonicity judgement in the experiments (see Table 3 & Table 8).[9]

The antonyms, synonyms and unrelated pairings that were retrieved on the basis of the sentential co-occurrence in the BNC and the 11 pairs already investigated in Herrmann et al's experiments were used as test items in the experiments. It was important for us to include synonyms in the experiment too since, like antonyms, their relation is based on both similarity and difference. The synonym relation is a similarity relation between the meanings of different forms, while the antonym relation, seen from the point of view of lexical opposition, is a relation of difference across both meanings and forms (Storjohann, 2009; see also M. L. Murphy, 2003, pp. 167–168, for a short discussion about the fuzzy distinction between antonyms and synonyms). However, this difference presupposes that the meanings represent opposite poles/parts on the same dimension. It is therefore a reasonable assumption that the participants would judge synonyms to be low-degree opposites that would be lower than that for antonyms but higher than for unrelated. Rubenstein and Goodenough (1965) carried out a study on synonyms using both written production and judgement tests. In the judgement experiments participants were asked to organize 65 word-pairs according to decreasing similarity of meaning and assign a value between 0.0 and 4.0 to each pair. Their results support a scale of similarity.

Our prediction that the seven antonym pairs that scored the highest in the corpus data as well as the highest scoring scalar adjective pair *beautiful – ugly* from Herrmann et al's study form a group of canonical antonyms was borne out. It was shown that *weak – strong*, *small – large*, *light – dark*, *narrow – wide*, *thin – thick*, *bad – good*, *slow – fast* and *ugly – beautiful* were judged by the participants to be examples of very good antonym pairs. They differ significantly from the rest of the antonym pairings in the judgement experiment. What was unexpected was that the scorings for the synonyms and the unrelated pairings did not differ significantly and that both categories had low standard deviations. We expected both categories to give rise to results with a high degree of variation in the judgements given by the participants. The standard deviations, however, show that the participants agreed to a large extent about the ratings for the canonical antonyms, the synonyms and the unrelated, while there was a great deal of disagreement regarding antonyms.

The judgement experiments also ruled out the fact that some of the pairings may be judged as better antonyms in the opposite order because no significant difference with respect to the order of the individual antonyms was found. The upshot of the judgement experiment is that there is a small set of canonical antonyms and a much larger set of antonyms, and there is a third set consisting of synonyms and unrelated pairs. The result of the judgement experiment reflects the figures for sentential co-occurrence in the BNC. The same pattern is also reflected in the response times. Moreover, regarding the 11 pairs that we included form Herrmann

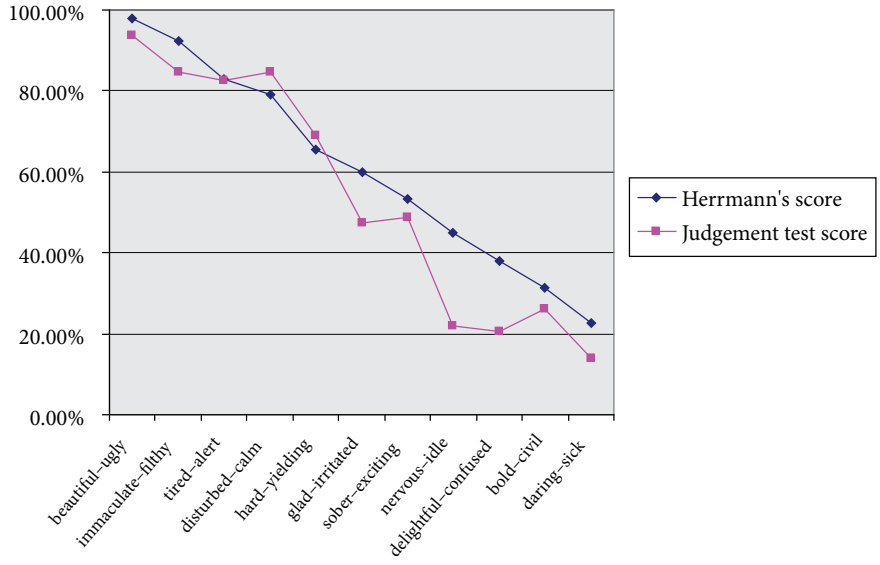

Figure 10. The scorings of Herrmann et al's (1986) word-pairs in relation to the present judgement test.

et al's (1986) study, the results of our judgement experiment was consistent with their result as shown in Figure 10 below.

Next, the pairs that we used as test items in judgement experiment were also used in our elicitation experiments. In that experiment the task of the participants was to provide the best antonym they could think of for a given seed word. The elicitation thus complements both the corpus search and the judgement experiment in that it was designed to tap the participants' memory for non-contextualized lexico-semantic knowledge. The overall prediction was that the elicitations would form a curve from total participant agreement for the strongly conventionalized antonyms to low participant agreement for the weakly conventionalized pairings. There were altogether 85 test items in the elicitation experiment — all the individual adjectives retrieved from the corpus and the selection from Herrmann et al's (1986) study. This means that the participants were also asked about antonyms for all the test items including the ones that were retrieved as synonyms and unrelated pairings from the corpus that were subsequently used as the synonym and the unrelated conditions in the judgement experiment. Comparisons of conventionalized antonym affinity can therefore only be carried out on the items that were included as pairs of antonyms in the judgement experiment.

The principal outcome of the elicitation experiment is that there are a limited number of test items that elicit one or two opposites only. All the participants suggested one and the same antonym for *bad* (*good*), *beautiful* (*ugly*), *clean* (*dirty*), *heavy* (*light*), *hot* (*cold*), *poor* (*rich*) and *weak* (*strong*) — yet, not in the opposite order, that is, not all participants supplied *bad* as an opposite of *good*. Thirteen test items yielded two antonyms: *black*, *fast*, *narrow*, *slow*, *soft*, *good*, *hard*, *open*, *big*, *easy*, *white*, *light*, *dark*, six elicited three antonyms: *large*, *rapid*, *small*, *ugly*, *exciting*, *thick* and so forth. The test item the highest number of antonyms (29 different antonyms, was *calm* (29) (see Appendix C for the full list).

The elicitation results for each of the test words are shown in Appendix C as a two-dimensional representation of the distribution of every elicitation across all the test words. Figure 6 shows the three-dimensional distribution of elicitations across all the test items and also the number of times each antonym was elicited. The top left-hand side of Figure 6 shows the seven test items for which all the participants suggested one and the same antonym and at the bottom right-hand side of the Figure 6 are the test items that gave rise to the largest number of different antonyms. There is a steady cline between the two extremes. There is also a cline within the scope of each test item from the antonyms that most participants considered the best antonym of *calm*, that is, *stressed* to the antonym that only one participant considered to be a good antonym of *calm*, that is, *troubled*.

The design of the experiment was such that there were no constraints on the elicitation. On the contrary, the purpose of the elicitation experiment, in contrast

to the judgement experiment, was to encourage the participants to respond spontaneously to the trigger word. This means that the participants were free to interpret the test items in any way and chose their antonyms accordingly. The pattern of the curve then is not only a reflection of a word's antonym partner in a certain context but also a reflection of their various readings in different contexts as well as the relative strength of these when no context is given. This reflects the potential of the adjectival test items to be construed as belonging to more than one semantic dimension due to the nature of the nominal meaning structure that they modify (Murphy & Andrew, 1993, Paradis, 2005).

To take the scale of MERIT as an example, as Figure 8 showed, forty-two out of the fifty participants suggested *bad* when presented with *good*, while only eight suggested *evil*. This result indicates a strong coupling between *good* and *bad*. It also shows that for the majority of the participants *good* is more strongly coupled with *bad* than with *evil*. This suggests the semantic dimension underlying the relation is more salient for most of the participants than the semantic dimension between *good* and *evil*. The pattern across the responses for *good* in contrast to responses for *bad* and *evil* is a simple pattern which places itself among the strongly conventionalized pairings. Figure 8 also showed that *mediocre – good* has a very weak relation.

Figure 11 shows a more complex map of relations involving *thick – thin*. *Thick* elicits *thin* in 45 cases, while *thin* only elicits *thick* in 13 cases. Instead, *thin* has a

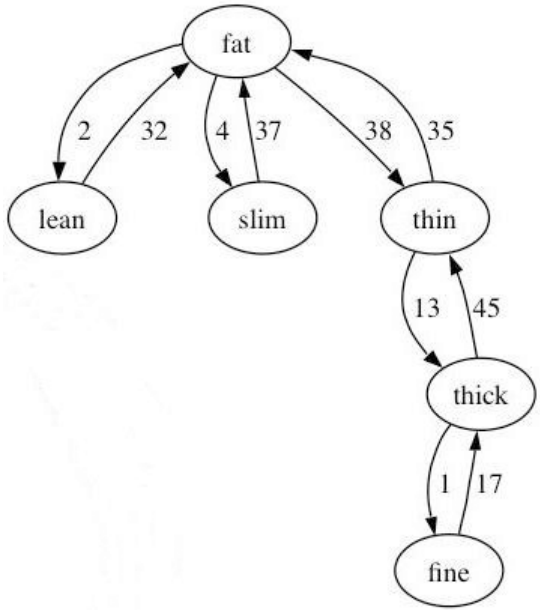

Figure 11. Relations between *fat*, *lean*, *slim*, *thin*, *thick*, and *fine* based on the elicitation experiment. The number of responses is marked by each arrow.

strong relation to *fat* and so does *fat* to *thin*. Both *lean* and *slim* have a relatively strong relation with *fat*, but this relation is very weak in the other direction. The reason for this is probably that *fat* has wider contextual application, while *lean* and *slim* are used in a more limited range of contexts. The same is true of the relationship between *thick* and *fine*, where *thick* has wider application than *fine*.

In relation to the number of antonyms elicited per test item, the number of times the participants abstained from providing an antonym is of interest. In the group of test items for which all the participants suggested antonyms, the mean of suggestions is 5.3 and the median 3, while for the group where participants neglected to suggest an antonym, the mean is 12.5 and the median is 13. This indicates that the items that have strongly conventionalized antonymic partners, that is, the canonical antonyms, are easier to retrieve from memory, while test items which are not members of conventionalized antonymic pairings are more demanding for the participants to identify, which results in more omitted responses as well as a larger number of different elicitations.

Comparing the results

A cluster analysis was carried out in order to see if all the pairings that were deemed canonical in the judgement experiment were also felt to be strong pairings in the elicitation experiment. The result of the cluster analysis is shown in Figure 9. The pairings fall into two main clusters. Both Cluster 1 and 2 contain partly different pairings from the ones in the judgement experiment. The reason for this is that all the individual words from the data bank were used as seed words in the elicitation. This means that some pairs that rank highly in the cluster analysis, for example, *big – small*, *black – white*, *fat – thin*, were not included as pairings in the judgement experiment, and for that reason no comparison can be made concerning them. In addition to our eight canonical pairings this list also contains other pairings that native speakers intuitively perceive as good pairings. None of the canonical pairings from the judgement experiment was found in the larger group consisting of 38 pairings.

On the surface it may look as if the corpus-driven retrieval of data for this study and the two different types of experiments yield slightly different results. The results of the corpus-driven extractions show clearly that there is a limited number of very strong pairings and a larger number of pairings which are strongly coupled but less so than the ones we called canonical antonyms. The same pairings were also judged to be the best pairings in the judgement experiment in which the participants were instructed to make conscious decisions about good and bad antonyms. The canonical antonyms were shown to be significantly different from

the rest of the antonym pairings, and these antonyms in turn were shown to differ significantly from synonyms and unrelated. In contrast to the results of the judgement experiments, the results of the elicitation experiment point to a cline from a very limited number of test items, that is, *bad*, *beautiful*, *clean*, *heavy*, *hot*, *poor* and *weak*, for which the participants were in total agreement about the best antonym of the test item *calm* for which the participants delivered 29 different suggestions. Among these 29 suggestions there was a continuum from stronger to very weak agreement across the participants in terms of how many suggestions each of the 29 antonyms elicited.

Nevertheless, the results of the elicitation experiment are not in conflict with the results from the judgement experiment. In the judgement experiment three significantly different groups crystallized (canonical antonyms, antonyms and the rest), while when the participants were given more freedom as in the elicitation experiment, the result was instead a cline with no clear cut-off points. The evidence from the elicitation experiment reveals the internal prototypicality structure of the category of antonymy, while the judgement experiments shows that there is in fact significant difference between canonical antonyms and other more contextual and less conventionalized pairings. The seemingly diverging results are of course partly due to the nature of the experiments as well as the design of the experiments and the statistical calculations. In the judgement experiment two distinct types were assumed as part of the design, while no such assumptions were involved for the elicitation experiment, neither in the design of the experiments nor in the statistical calculations. The corpus-driven extraction method was not associated with minimum assumptions besides the actual dimensions selected for the scope of this study.

We thus analyze the results from the different parts of this study as pointing in the same direction, namely that at the level of lexico-semantic relations there are few strongly associated pairs and there is also a large group of pairings that are less strongly associated, at least at the lexical level and with minimum ontological context. We interpret these seemingly contradictory results as an indication of a core canonical antonyms and a large number of pairings from more to less strongly lexically coupled. Granted that we take the patterning of the results of this study to reveal something about the category of antonymy in language, the picture that emerges is a prototypicality structure with a centre and a steadily fading strength of conventionalized opposability and so it proceeds *ad infinitum*.

Again, these results reflect our corpus-driven method, preceding the selection of test items for the test set, and the two different experiment types. In another study, Jones, Paradis, Murphy, and Willners (2007) we used the World-Wide-Web as corpus. We approached the issue of antonym canonicity by building specifically on research that has demonstrated the tendency of antonyms to favour certain

lexico-grammatical constructions in discourse, such as *X and Y alike*, *between X and Y*, *both X and Y*, *either X or Y*, *from X to Y*, *X versus Y* and *whether X or Y*, as identified by Jones (2002). That study argued that canonical antonyms can be expected to co-occur with high fidelity in such constructions. Fourteen contrastive constructions were used for retrieval of a range of contrast items across a number of seed words. Strong correlations emerged between those items retrieved most frequently and adjectives cited as 'good opposites' in the elicitation experiments. Indeed, in the case of nine of the ten seed words selected as a starting point for the web searches, that is, *beautiful*, *poor*, *open*, *large*, *rapid*, *exciting*, *strong*, *wide*, *thin* and *dull*, the adjectives retrieved most often in searches were the same as the adjectives that were suggested by the participants in the elicitation experiment. Only *thin* did not retrieve *fat* most frequently in the web study, but the antonym that was retrieved most commonly was instead *thick* which ranked second in the elicitation experiment (cf. Figure 11). However, there is agreement between the corpus-driven ranking and the judgement experiment in the present study where *thick* was found to have the strongest co-occurrence rates. Unfortunately, the *thin – fat* pairing was not a test item in the judgement experiment.

A striking result from the web study was that the second most reliable antonym of *open* is *laparoscopic*, describing two types of surgery. The strength of the coupling of *open – laparoscopic* in terms of co-occurrence within a large proportion of antonymic frames makes them a strong candidate to be considered a canonical pair. At the same time, it could be stated with some confidence that English speakers would be very unlikely to propose *laparoscopic* as an antonym of *open* in an elicitation test. However, those lay people who know the meaning of *laparoscopic* would probably suggest *open* as antonym. Also, *open – laparoscopic* would be likely to be considered a pair of good antonyms in a judgement experiment. The upshot of this is that the strength of pairings in more restricted registers and genres in text are not necessarily the antonyms favoured in experiments. Experimental elicitation reflects associative strength, frequency and contextual versatility because participants offer well-known opposites of salient semantic dimensions. The less contextually constrained the pairings are, the more strongly they will elicit one another in context-free elicitation experiments. Goodness of opposability and antonym affinity can not be reduced to mere word/sense frequency. Still, frequency should not be altogether dismissed because the strength of affinity of pairs such as *open – laparoscopic* is most likely a matter of frequency in the context of surgery. Again, this demonstrates that different techniques yield slightly different results for a complex and context-dependent relation such as antonymy.

Conclusion

The aim of this paper was to investigate the nature of antonymy in language using scalar adjectives in English as test items in order to find out whether there are good antonym pairings, less good pairings and pairings that are not antonymous at all. In addition to that question there are two corollary questions, namely what the characteristics of the more strongly conventionalized canonical pairings are, and what the characteristics of the less closely associated word pairs are. An important secondary aim of the study was to carry out the investigation using different methods in order to see how the results were influenced by the different techniques. For that purpose we were using both textual and psycholinguistic methods. The textual method was mainly used to generate data, but in itself it also yielded results. The extraction of the test items for the psycholinguistic experiments was carried out using a computer program *Coco* for retrieval of antonyms, synonyms and unrelated pairings along seven dimensions of meaning, namely SPEED, LUMINOSITY, STRENGTH, SIZE, WIDTH, MERIT and THICKNESS primarily represented by the pairings *slow – fast*, *light – dark*, *weak – strong*, *small – large*, *narrow – wide*, *bad – good* and *thin – thick* respectively. We also included a small subset of test items from a study by Herrmann et al. (1986) of which the strongest pairing was *beautiful – ugly*. These test items (and combinations of their various synonyms as pairs of antonyms, synonyms and unrelated significantly co-occurring in sentences in the BNC) were used as the main body of test items for the experiments. Two types of experiments were carried out: a judgement experiment in which the participants were asked to evaluate the goodness of a given word pair on a scale from very bad opposites to excellent opposites, and an elicitation experiment in which the participants were given the complete list of the same test items and were asked to provide the best antonym for each of the individual seed words.

The hypothesis under investigation was that a scale exists between, at the one extreme, pairings that are strongly conventionalized as antonyms and, at the other extreme, pairings which may be opposable in some context and pairings for which it is very difficult to think of a context in which they could be used as antonyms. We also hypothesized that there will be a very limited number of canonical antonyms that would elicit complete or almost complete agreement across the participants about their special status as canonical antonyms. These pairings would coincide with the pairings retrieved from the corpus with the strongest figures for sentential co-occurrence, that is, the above-mentioned pairs that are the principal representatives of the meaning dimensions. Both hypotheses were proven correct with some qualification. The corpus-driven method of extraction showed that the above antonym pairings co-occurred sententially more strongly than any of the other pairings that were used in the experiments (see Appendix A). The results

from the judgement experiment revealed a significant difference between the above eight pairings and the rest of the antonym pairings as well as a difference between the antonym pairings and the synonym pairings. There was no significant difference between the synonyms and the unrelated pairings.

The investigation showed that, on the one hand there is a small but distinct group of conventionalized canonical antonyms. This was shown mainly by the outcome of the judgement experiment since the design and the statistical calculations of that experiment were geared towards the boundaries between the four conditions, that is, canonical antonyms, antonyms, synonyms and unrelated. On the other hand, the elicitation experiment points up a continuum from excellent antonym pairings with a total participant consensus to pairings with a steady decrease in agreement. The elicitation experiment and the attendant calculations were designed to shed light on the structure rather than on the borders. The outcome of both the experiments and the co-occurrence statistics converge in a picture of the category of antonymic lexical meanings in English as a prototypicality structure with a small number of excellent representatives of the category, to category members on the outskirts that are hard to conceive of as antonym pairings (Paradis, 2009; Paradis & Willners, submitted).

We conclude that the lexical items that are most appropriate to be considered canonical all co-occur frequently in the BNC. They top the list both taken individually and in pairs in the corpus study. The actual frequency for adjectival meanings of this type is taken to be a sign of their being applicable in large range of meaning structures and useful in a large range of contexts individually and as pairs in which case the dimensional meaning structure is guaranteed. The relationship between the members of the canonical pairings is also symmetrical in the majority of the cases. This means that sequential order does not seem to play any role in the judgement of goodness and in the case of elicitations both members trigger the other to the same extent. Another factor that seems to be of importance for the best pairings, judging from the experiment results, is the salience of the dimension. The dimension of which the canonical antonyms are representatives is salient in the sense that it is easily identifiable. For instance, the SPEED dimension underlying *slow – fast* is easily identifiable, while the dimension behind say *rare – abundant*, *calm – disturbed*, *lean – fat*, *open – laparoscopic* or *narrow – open* are not. This has to do with the more specialized ontological applications of these adjectives to nominal meanings which concern different readings and sometimes also different meanings of these words and to certain very restricted styles and genres. This also means that polysemy and multiple readings as such do not prevent a word from participating in a canonical relation with another word. For instance, *light – dark* and *light – heavy*, *narrow – wide* and *narrow – open*. Contextual versatility is a reflection of ontological versatility, that is, that the use potential of these antonyms

applies in a wide range of ontological domains, and they are frequent in constructions and contrasting frames in text and discourse.

Two approaches to antonymy were set up as contrasting positions: the *lexical categorical* approach and the *cognitive prototype continuum* approach. The position taken by proponents of the former approach is that antonymy is a lexical relation and words are either lexical antonyms or not. Antonyms are pre-stored and get their meanings from the relation of which they form part. The model is static and context insensitive. Words either have antonyms or not. If they have antonyms they have one antonym. For instance, Miller and Fellbaum (1991, p. 210) state that *ponderous* is often used where *heavy* would also be felicitous, but unlike *heavy* it has no antonym. Similarly *heavy* and *weighty* have very similar meanings but different antonyms, *light* and *weightless* respectively. If antonymy was a conceptual relation, people would have accepted *weighty* and *light* or *heavy* and *weightless* as pairs of antonyms, which thus is not the case according to the authors. The conceptual opposition in their model between, say, *ponderous* and *light* is mediated by *heavy*. Conceptual opposition is an effect of lexical relations rather than a cause. Our experiments paint a totally different picture. It is obvious, in particular from the elicitation experiment, that the participants have very different scenarios and different styles and genres in mind, when they offer antonyms to adjectives. The lexical categorical approach has no explanations for these patterns. Also, they predict a definite boundary between adjectives such as *heavy* that have antonyms and adjectives such as *ponderous* that have no antonyms on grounds that are not empirically supported. The prediction that falls from such a position is that we would obtain high scores which are consistent across native speakers for all adjectives that have antonyms and no responses for words with no antonyms, such as *ponderous*. In the lexical categorical model, antonymy as a category will be monolithic without any internal structure

The cognitive prototype approach, on the other hand, takes antonymy to have conceptual basis. Antonymy is a construal rather than a pre-stored representation. It is dependant on general cognitive processes such as comparison and attention and relies on a binary configuration of a segment of content (Paradis, 2009; Paradis & Willners, submitted). Adjectival meanings are fostered in conceptual combinations with nominal meanings. Conceptual structures are the cause of antonym couplings, not an effect, and salient contentful dimensions such as SPEED, LUMINOSITY, STRENGTH, SIZE, WIDTH, MERIT and THICKNESS form good breeding grounds for routinization of lexical pairings (Herrmann et al., 1986). This approach predicts a category with an inherent continuum structure with a small number of core members associated with particularly salient dimensions. The results of this investigation indicate that such canonical pairing have lexical correlates, while the vast majority of antonyms have only associatively weak partners in

situations when speakers are invited to produce or evaluate antonyms without any contextual constraints. Given a specific context, antonym couplings are bound to be stronger and more consistent across speakers (Murphy & Andrew, 1993). In the lexical categorical model different contexts do not affect the antonym, since the antonym of a word is not determined by the context and sense. Finally, the prototype continuum model is consistent with categorization in general (Taylor, 2003).

The theoretical implication of our investigation is that antonymy is primarily a conceptual relation in that binary contrast is always a possibility in meaning construals and such construals are based on general knowledge-intensive cognitive processes. However, in spite of the fact that other antonyms would be possible, our study also indicates that a select group of antonyms are lexically recognized and particularly strongly associated in memory. For instance, not much imagination is required to produce possible antonyms of *bad* (*satisfactory*, *beneficial*, *fine*, *obedient*), but, nevertheless, all of the experiment participants suggested *good*. Pairings for which the participants suggested many different antonyms in the elicitation experiment are more likely to be contextually idiosyncratic, that is, not strongly routinized as pairs in our minds, or very weakly conventionalized, more generally, due to extreme genre or register restrictions.

Finally, while binary contrast is generally regarded as an extremely powerful construal in human thinking and in associative strength between words, research on language and cognition calls for evidence from different sources and cross-fertilization of scientific techniques. Our next step will be extended studies using corpus methodologies as well as both psycholinguistic and neurolinguistic methods in order to shed more light on contrast in language, thought and memory.

Acknowledgement

We gratefully acknowledge detailed comments from Lena Ekberg, Lynne Murphy and two anonymous reviewers. We wish to express particular gratitude Joost van de Weijer for help with the statistics, Anders Sjöström for help with the figures and to Simone Löhndorf and Lynne Murphy for help with data collection. We also wish to extend our thanks to Gary Libben, Jordan Zlatev and the semantics seminar at Lund University.

Notes

1. It should be noted that we use *antonymy* as a cover term for all different kinds of oppositeness in this paper. This is different from how the term is used in most of the literature, for example in Croft and Cruse (2004), Cruse (1986), Lyons (1977) and Paradis (1997) where antonymy is reserved for opposites that are associated with a scale. For various definitions and studies of

antonymy see Cruse (1986), Fellbaum (1998), Jones (2002), M. L. Murphy (2003) and http://www.f.waseda.jp/vicky/complexica/, Lehrer and Lehrer (1982), Muehleisen (1997), Paradis (1997, 2001).

2. Exceptions are Gries and Otani (in press) and Willners and Paradis (in press). Otherwise, there is a number of corpus studies: Muehleisen (1997), Jones (2002, 2006, 2007), Jones & Murphy (2005), Jones et al. (2007), Murphy & Jones (2008), Murphy et al. (2009), Muehleisen & Otani (2009), Storjohann (2009), Tribushinina (in preparation), Willners (2001), and experimental studies: Becker (1980), Deese (1965), Herrmann et al. (1979), Holleman & Pander Maat (2009), Paradis & Willners (2006, in preparation).

3. It deserves to be pointed out already here that X and Y are to be considered as two different variables in Table 2 and 3. They are not given in any particular order. In Table 6, however, the words within the pairs are ordered on semantic grounds to match the design of the experiment.

4. There are also six other pairs (both antonyms, synonyms and unrelated) from the test set in Appendix A (*fast – rapid*, *small – tiny*, *evil – good*, *bad – poor*, *good – healthy*, *heavy – thick*).

5. We made a web search for the ordering of the canonical pairings and found that the order X, Y was the more common in the frames "X and Y" and "X or Y" across all seven dimensions. There are of course other frames, but the conjunctions were used because they are also common in irreversible binomials, for example, *fast and loose*, *loud and clear*, *sick and tired*, *high and dry*, *thick and thin*, *sweet and sour*, *neat and tidy*, *peace and quiet*.

6. For more information about E-prime on http://www.pstnet.com/products/e-prime/.

7. By mistake, the test item *gloomy-bright* appeared as *gloomy-gloomy* on the screen. The word pair was excluded from the analysis leaving us with 52 test items in total.

8. This does not mean that items that are less frequent in language cannot form strongly conventionalized canonical pairings. For instance, had we included verbs, it is most likely that *maximize – minimize* would have scored high both in terms of sentential co-occurrence and in the experimental investigations, as indeed was shown by Herrmann et al. (1986). The same was shown by Jones et al.'s (2007) web study of antonyms in constructions as described in the discussion section.

9. In current empirical research where corpora are used, a distinction is being made between corpus-based and corpus-driven methodologies (Francis, 1993; Paradis & Willners, 2007; Storjohann, 2005; Tognini-Bonelli, 2001; pp. 65–100). The distinction is that the corpus-based methodology makes use of the corpus to test hypotheses, expound theories or for retrieval of real examples, while in corpus-driven methodologies the corpus as such serves as the empirical basis from which researchers extract their data with minimum prior assumption. In the latter approach all claims are made on the basis of the corpus evidence with the necessary proviso that the researcher determines the search items in the first place. Our method is of a two-step type in that we mined the whole corpus for both individual occurrences and co-occurrence frequencies for all adjectives without any restrictions and from those data we selected our seven dimensions and all their synonyms.

References

Becker, C. A. (1980). Semantic context effects in visual word recognition: an analysis of semantic strategies. *Memory and Cognition, 8*, 493–512.

Charles, W. G., & Miller, G. A. (1989). Contexts of antonymous adjectives. *Applied Psycholinguistics, 10*, 357–375.

Charles, W. G., Reed, M. A., & Derryberry, D. (1994). Conceptual and associative processing in antonymy and synonymy. *Applied Psycholinguistics, 15*, 329–354.

Croft, W., & Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cruse, D. A. (1986). *Lexical semantics.* Cambridge: Cambridge University Press.

Cruse, D. A. (1994). Prototype theory and lexical relations. *Rivista di Linguistica, 6*,167–188.

Cruse, D. A. (2002). The construal of sense boundaries. *Revue de Sémantique et Pragmatique, 12*, 101–119

Deese, J. (1965). *The structure of associations in language and thought*. Baltimore: Johns Hopkins University Press.

Divjak, D., & Gries, S. (2008). Clusters in the mind? Converging evidence from near synonymy in Russian. *The Mental Lexicon, 3*(2), 188–213.

Fellbaum, C. (1995). Co-occurrence and antonymy. *International Journal of Lexicography, 8*, 281–303.

Fellbaum, C. (1998). *WordNet: An electronic lexical database.* Cambridge, MA: MIT Press.

Francis, G. (1993). A corpus-driven approach to grammar: principles methods and examples. In M. Baker, G. Francis, & E. Bonini-Tognelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 137–156). Amsterdam: John Benjamins.

Glynn, D., Geeraerts, D., & Speelman, D. (2007). *Entrenchment and social cognition. A corpus-driven study in the methodlogy of language description*. Paper presented at the first SALC (Swedish Association for Language and Cognition) conference at Lund University November 29-December 1, 2007.

Gries, S. Th., & Otani, N. (in press). Behavioral profiles: a corpus-based perspective on synonymy and antonymy. *ICAME Journal.*

Gries, S. Th., & Divjak, D. S. (in press). Behavioral profiles: a corpus-based approach towards cognitive semantic analysis. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics*. Amsterdam: John Benjamins.

Gross, D., Fischer U., & Miller, G. A. (1989). The organization of adjectival meanings. *Journal of Memory and Language, 28*, 92–106.

Gross, D., & Miller, K. J. (1990). Adjectives in WordNet. In WordNet: an on-line lexical database. *International Journal of Lexicography, 3*(4), 265–277.

Herrmann, D. J., Chaffin, R., Conti, G., Peters, D., & Robbins, P. H. (1979). Comprehension of antonymy and the generality of categorization models. *Journal of experimental psychology: human learning and memory, 5*, 585–597.

Herrmann, D. J., Chaffin, R., Daniel, M. P., & Wool, R. S. (1986). The role of elements of relation definition in antonym and synonym comprehension. *Zeitschrift für Psychologie, 194*, 133–153.

Holleman, B. C., & Pander Maat H. L. W. (2009). The pragmatics of profiling: Framing effects in text interpretation and text production. *Journal of Pragmatics, 41*(11), 2204–2221.

Jones, S. (2002). *Antonymy: a corpus-based perspective*. London: Routledge.

Jones, S. (2006). The discourse functions of antonymy in spoken English. *Text and Talk, 26*(1), 191–216.
Jones, S. (2007). 'Opposites' in discourse: a comparison of antonym use across four domains. *Journal of Pragmatics, 39*(6), 1105–1119.
Jones, S., & Murphy, M. L. (2005). Using corpora to investigate antonym acquisition. *International Journal of Corpus Linguistics, 10*(3), 401–422.
Jones, S., Paradis, C., Murphy, M. L. & Willners, C. (2007). Googling for opposites: a web-based study of antonym canonicity. *Corpora, 2*(2), 129–154.
Justeson, J. S., & Katz, S. M. (1991). Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistic,s 17*, 1–19.
Justeson, J. S., & Katz, S. M. (1992). Redefining antonymy: the textual structure of a semantic relation. *Literary and Linguistic Computing, 7*, 176–184.
Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: Stanford University Press.
Langacker, R. W. (1991). *Grammar and concptualization*. Berlin: Mouton de Gruyter.
Lehrer, A. (1985). Markedness and antonymy. *Journal of Linguistics, 21*, 397–429.
Lehrer, A. (2002). Gradable antonymy and complementarity. In A. D. Cruse, F. Hundsnurcher, M. Job, & P. R. Lutzeier (Eds.), *Handbook of lexicology* (pp. 498–508). Berlin: Mouton de Gruyter.
Lehrer, A., & Lehrer, K. (1982). Antonymy. *Linguistics and Philosophy, 5*, 483–501.
Libben, G., & Jarema, G. (2002). Mental lexicon research in the new millennium. *Brain and Language, 81*, 2–11.
Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
Mettinger, A. (1994). *Aspects of semantic opposition*. Oxford: Clarendon Press.
Mettinger, A. (1999). *Contrastivity: the interplay of language and mind*. Unpublished Habitilationsschrift, University of Vienna.
Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition, 41*, 197–229.
Muehleisen, V. (1997). *Antonymy and Semantic Range in English*. PhD. dissertation, Evanston, IL, Northwest University.
Muehleisen, V., & Isono, M. (2009). Antonymous adjectives in Japanese discourse. *Journal of Pragmatics, 41*(11), 2185–2203.
Murphy, G. L., & Andrew, J. M. (1993). The conceptual basis of antonymy and synonymy in adjectives. *Journal of Memory and Language, 32*, 301–319.
Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press.
Murphy, M. L. (2003). *Semantic relations and the lexicon: antonyms, synonyms and other semantic paradigms*. Cambridge: Cambridge University Press.
Murphy, M. L. (2006). Antonyms as lexical constructions: or, why paradigmatic construction is not an oxymoron. *Constructions* SV1–8/2006 (Available at http://www.constructions-online.de).
Murphy, M. L., & Jones, S. (2008). Antonymy in children's and child directed speech. *First Language, 28*(4), 403–30.
Murphy, M. L., Paradis, C., Willners C., & Jones, S. (2009). Discourse functions of antonymy: a cross-linguistic investigation. *Journal of Pragmatics, 41*(11), 2159–2184.
Ogden, C. K. (1932). *Opposition: A linguistic and psychological analysis*. Bloomington: Indiana University Press.
Osgood, C., Suci, George J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Palermo, D. S., & Jenkins, J. J. (1964). *Word association norms: grade school through college*. Minneapolis: University of Minnesota Press.
Paradis, C. (1997). *Degree modifiers of adjectives in spoken British English*. Lund: Lund University Press.
Paradis, C. (2001). Adjectives and boundedness. *Cognitive Linguistics, 12*(1), 47–65.
Paradis, C. (2003). Is the notion of linguistic competence relevant in Cognitive Linguistics? *Annual Review of Cognitive Linguistics, 1*, 247–271.
Paradis, C. (2005). Ontologies and construal in lexical semantics. *Axiomathes, 15*, 541–573.
Paradis, C. (2008). Configurations, construals and change: expressions of degree. *English Language and Linguistics,12*(2), 317–343.
Paradis, C. (2009). A dynamic construal approach to antonymy. *Proceedings the 19th symposium for theoretical and applied linguistics April 3–5, 2009*.
Paradis, C., & Willners, C. (2006). Antonymy and negation: the boundedness hypothesis. *Journal of Pragmatics, 38*(7), 1051–1080.
Paradis, C., & Willners, C. (2007). Antonyms in dictionary entries: methodological aspects. *Studia Linguistica, 61*(3), 261–277.
Paradis, C. & Willners, C (2009). Negation and approximation as configuration construal in SPACE. In C. Paradis, J. Hudson & U. Magnusson (Eds.), *Conceptual Spaces and the Construal of Spatial Meaning. Empirical evidence from human communication* (in preparation). Unpublished.
Paradis, C. & Willners, C. (submitted). Antonymy: from conventionalization to meaning-making.
Rubenstein, H., & Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM, 8*(10), 627–633.
Sampson, G. (2000). Review of WordNet: an electronic lexical database. *International Journal of Lexicography, 13*, 54–59.
Storjohann, P. (2005). Corpus-driven vs. corpus-based approach to the study of relational patterns. *Proceedings of the Corpus Linguistics conference 2005 in Birmingham, 1*(1), http://www.corpus.bham.ac.uk/pclc/index.shtml#complie.
Storjohann, P. (2009) Plesionymy: A case of synonymy or contrast? *Journal of Pragmatics, 41*(11), 2140–2158.
Talmy, L. (2000). *Towards a cognitive semantics*. Cambridge (MA): The MIT Press.
Taylor (2003). *Linguistic Categorization*. Oxford: Oxford University Press.
Tognini-Bonelli, E. (2001). *Corpus Linguistics at work. Studies in Corpus Linguistics 6.* Amsterdam/Philadelphia: John Benjamins.
Tomasello, M. (2003). *Constructing a language*: *a usage-based theory of language acquisition*. Cambridge (MA): Harvard University Press.
Tribushinina, E. (2009). Spatial adjectives in Dutch child language: towards a usage-based model of adjective acquisition. In C. Paradis, J. Hudson, & U. Magnusson (Eds.), *Conceptual Spaces and the Construal of Spatial Meaning. Empirical evidence from human communication* (in preparation). Unpublished.
Verhagen (2005) *Constructions of intersubjectivity: discourse, syntax and cognition*. Oxford: Oxford University Press.
Willners, C. (2001). Antonyms in Context. A Corpus-based Semantic Analysis of Swedish Descriptive Adjective. *Working Papers, Department of Linguistics, Lund University. 40*.

Willners, C., & Holtsberg A. (2001). Statistics for sentential co-occurrence. *Working Papers, Department of Linguistics, Lund University*, *48*, 135–148.

Willners, C., & Paradis, C. (in press) Swedish opposites — a multi-method approach to antonym canonicity. In P. Storjohann (Ed.), *Lexical-semantic relations from theoretical and practical perspectives. Lingvisticæ Investigationes Supplementa*, John Benjamins, Amsterdam.

Appendix A

The Ten Most Frequently Co-occurring Word Pairs from each Dimension in the Study (Comprising more than 68,000 Pairs) Sorted According to Falling Frequency of Co-occurrence within each Dimension

$Word_X$	$Word_Y$	N_X	N_Y	Co
slow	**fast**	5760	6707	163
rapid	slow	3526	5760	54
quick	slow	6670	5760	39
firm	smooth	6157	3052	34
fast	quick	6707	6670	34
fast	rapid	6707	3526	29
hot	sudden	9445	3920	23
gradual	sudden	1066	3920	22
gradual	slow	1066	5760	22
hot	smooth	9445	3052	19
black	white	19998	19184	2663
blue	white	10157	19184	753
dark	white	12907	19184	530
black	dark	19998	12907	447
blue	dark	10157	12907	431
light	**dark**	12396	12907	402
black	blue	19998	10157	359
blue	pale	10157	3807	305
dark	pale	12907	3807	236
blue	bright	10157	6181	231
weak	**strong**	4522	19550	455
hard	soft	18212	6626	270
heavy	light	10537	12396	233
soft	warm	6626	7039	154
powerful	strong	7213	19550	138
light	strong	12396	19550	135

Appendix A. *(continued)*

$\mathbf{Word_X}$	$\mathbf{Word_Y}$	$\mathbf{N_X}$	$\mathbf{N_Y}$	**Co**
hard	strong	18212	19550	120
heavy	strong	10537	19550	110
light	warm	12396	7039	102
pale	thin	3807	5536	89
small	**large**	51865	47184	3642
big	small	33688	51865	844
low	small	28903	51865	656
large	low	47184	28903	611
big	large	33688	47184	462
large	wide	47184	16812	387
small	wide	51865	16812	342
heavy	large	10537	47184	267
small	tiny	51865	5570	246
important	wide	39264	16812	229
big	large	33688	47184	462
large	wide	47184	16812	387
large	open	47184	19320	327
heavy	large	10537	47184	267
open	wide	19320	16812	236
narrow	**wide**	5338	16812	191
deep	wide	9817	16812	152
broad	large	6450	47184	152
narrow	broad	5338	6450	140
big	heavy	33688	10537	139
bad	**good**	26204	124542	1957
evil	good	2291	124542	300
good	pretty	124542	3924	265
bad	poor	26204	16579	217
big	white	33688	19184	217
good	healthy	124542	3970	208
fine	white	15331	19184	141
good	sound	124542	2513	135
complete	full	9666	28529	119
great	moral	62347	5118	119
heavy	light	10537	12396	233
narrow	wide	5338	16812	191
deep	wide	9817	16812	152

Appendix A. *(continued)*

$Word_X$	$Word_Y$	N_X	N_Y	Co
broad	narrow	6450	5338	140
thin	**thick**	5536	5119	130
broad	wide	6450	16812	113
heavy	thick	10537	5119	105
pale	thin	3807	5536	89
fat	thin	3903	5536	78
deep	light	9817	12396	73

Appendix B. Instructions for Elicitation Experiment

You are going to be given a list with 85 English words. For each word, write down the word that you think is the best opposite for it in the blank line next to it.

- Don't think too hard about it — write the first opposite that you think of.
- There are no 'wrong' answers.
- Give only one answer for each word.
- Give opposites for all the words, even when the word doesn't seem to have an obvious opposite
- Don't use the word *not* in order to create an opposite phrase. Your answer should be one word.

Example: The opposite of MASCULINE is _____

You might answer *feminine*.
You may leave the experiment whenever you want to. Your answers are anonymous and will be handled confidentially. If you are interested in knowing about the aims or results of this experiment or if you have any other questions, please let us know.

Appendix C. Stimuli and Responses in the Elicitation Experiment

Stimuli in bold followed by the responses for each stimulus ordered according to falling frequency. The stimuli are ordered according to rising number of responses. Omitted responses are not included.

bad good
beautiful ugly
clean dirty
heavy light
hot cold
poor rich
weak strong

young old
black white colour
fast slow fast
narrow wide broad
slow fast quick
soft hard rough
good bad evil
hard soft easy
open closed shut
big small little
easy hard difficult
white black dark
light dark heavy
dark light pale
large small little slim
rapid slow sluggish fast
small big large tall
ugly beautiful pretty attractive
exciting boring dull unexciting
thick thin clever fine
strong weak feeble mild slight
wide narrow thin skinny slim
evil good kind angelic pure
thin fat thick overweight wide
sober drunk frivolous inebriated intoxicated pissed
filthy clean spotless immaculate pristine sparkling
huge tiny small little minute petite
sick well healthy fine ill yum
enormous tiny miniscule small little minute slight
dull bright exciting interesting shiny lively sharp
bright dark dull dim gloomy stupid obscure
fat thin slim Iean skinny thick wrong
rare common comonplace ubiquitous frequent plentiful well-known
feeble strong robust hard impressive powerful steadfast
broad narrow thin slim small lean slight
smooth rough bumpy hard jagged hairy resistent
healthy unhealthy sick ill lame diseased poorly sickly
tiny huge large big enormous massive giant gigantic
lean fat fatty flabby large plump support stocky wide
heroic cowardly unheroic scared wimpish villainous disappointing reticent weak
glad sad unhappy sorry upset disappointed regretful cross worried
bare covered clothed dressed abundant cluttered full loaded patterned
slim fat broad big chubby wide large obese plump round
tough weak tender easy soft flimsy gentle sensitive weedy wimpy
gradual immediate sudden rapid fast quickly instant abrupt incremental swift
tired awake energetic alert lively fresh wakeful energized peppy perky rested

sudden gradual slow prolonged expected incremental immediate delayed foreseen infrequent predictable
idle busy active energetic hard-working working awake conscious diligent industrious proactive workaholic
gloomy bright happy cheerful cheery light sunny clear illumined nice merry pleasant
tender tough rough hard well-done cold robust chewy harsh mean nash strong uncaring
pale dark bright tanned bold brown coloured red ruddy colourful healthy rosey swarthy vivid
nervous calm confident bold brave relaxed alert assured excited fine innervous ready steady uncaring
limited unlimited extensive abundent comprehensive endless plenty available broad capacious common fat infinite widespread
robust weak fragile feeble flimsy shoddy thin brittle frail lethagic natural skinny slim vunerable
fine thick coarse bad bold dull wide blunt clumsy cloudy mad ok rough wet unwell
abundant scarce rare sparse little lacking disciplined few limited needed none meagre plentiful sparing threadbare
pure impure tainted contaminated corrupt dirty tarnished evil adulterated bad foul mixture sinful unclean unpure
immaculate untidy dirty messy filthy scruffy dishevelled boring faulty ramshacky spotted stained tarnished terrible tawdry
civil uncivil rude anarchic barbaric belligerent childish corperate couth horrible impolite mean nasty military savage unfair
extensive limited small intensive narrow restricted brief minimal constrained inextensive insufficient scanty short superficial sparse unextensive vague
grim nice happy bright cheerful pleasant positive hopeful good pleasant carefree clear cosy fun jolly reassuring welcoming
slender fat broad plump wide bulky chubby thick well-built big chunky curvy lean massive obese podgy portly rotund
delicate robust strong tough sturdy hardy rough coarse unbreakable bold bulky crude course gross hard hard-wearing harsh heavy
immediate later delayed slow gradual distant deferred extended anon eventually far forever longterm pending postponed prolonged soon whenever
modest boastful immodest arrogant bigheaded brash conceited extravagant vain outgoing blasé confident forward ignorant modest proud quiet shy
great small rubbish terrible bad average awful crap dreadful insignificant lowlyI mediocre microscopic obscure ok shit tiny poor unremarkable
firm soft weak floppy lenient wobbly flexible flimsy gentle groundless relenting lax limp loose saggy shaky undecided unsolid unstable
confused clear understood knowing sure lucid organised certain alert clued-up clearheaded coherent comprehending confident enlightened fine focused notconfused scatty together
bold timid shy cowardly faint fine italic thin nervous cautious faded feint frightened hairy meek quiet scared timorous weak yellow
daring cowardly timid nervous scared boring carefully cautious shy afraid careful fat faltering fearful reticient safe staid undaring wimpish
mediocre outstanding excellent exceptional brilliant amazing good great challenging charge clever extreme fair interesting mediocre rare special superb unusual wicked

yielding unyielding firm resisting dormant hard stubborn agressive dying fighting fixed lose losing obdurate rigid steadfast steamrollering strong stuck tough unproductive
irritated calm content relaxed amused fine placid serene soothed comfortable easy even good-humoured happy laid-back normal ok patient pleased tranquil unperturbed unruffled
alert sleepy tired asleep dozy oblivious distracted dull drowsy groggy lazy slow apathetic awake complacent dim dopey lethargic spacey torpid unaware unconscious unresponsive
disturbed calm undisturbed sane peaceful settled stable untouched alone balanced content fine ignored normal quiet relaxed together tranquil unaffected uninterrupted untroubled welcome well-adjusted well-balanced
slight large great strong big heavy considerable enormous huge major substantial very alot extensive heavyset lots marked massive plenty pronounced robust rough severe thick unslight well-built wide
delightful horrible awful unpleasant boring disgusting repulsive tedious abhorrent annoying crap difficult distasteful dredful dull grim hateful horrendous horrid irritating miserable nasty repellent revolting rubbish terrible uninteresting yuk
calm stressed stormy rough agitated excited hyper panicked angry annoyed anxious choppy crazy flustered frantic frenzied hectic hubbub hysterical irrate irrational jumpy lively loud nervous neurotic rage reckless tense troubled

Author's address

Carita Paradis
The School of Humanities
Växjö University
SE-351 95 Växjö
Sweden.

Carita.Paradis@vxu.se