



LUND UNIVERSITY

As lexical as it gets: The role of co-occurrence of antonyms in a visual lexical decision experiment

van de Weijer, Joost; Paradis, Carita; Willners, Caroline; Lindgren, Magnus

Published in:
Frequency effects in language representation

2012

[Link to publication](#)

Citation for published version (APA):
van de Weijer, J., Paradis, C., Willners, C., & Lindgren, M. (2012). As lexical as it gets: The role of co-occurrence of antonyms in a visual lexical decision experiment. In D. Divjak, & S. T. Gries (Eds.), *Frequency effects in language representation* (Vol. 2, pp. 255-279). Mouton de Gruyter.

Total number of authors:
4

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

As lexical as it gets the role of co-occurrence of antonyms in a visual lexical decision experiment*

Joost van de Weijer, Carita Paradis, Caroline Willners and Magnus
Lindgren
Lund University, Sweden

1. Introduction

When asked about their opinion of how good a pair of lexical items are as antonyms, speakers judge *slow-fast* to be a pair of strongly antonymic adjectives, while *slow-rapid*, *slow-express* and *slow-blistering* are perceived as less good pairings and *fast-dull* as less good pairing than all the others. This raises the question *why* some pairs are considered better pairings than others and *why* they form entrenched pairings in memory and conventionalized couplings in text. There are most likely a number of converging reasons for goodness of antonym pairings (*antonym canonicity*), i.e. the extent to which antonyms are both semantically related and conventionalized as pairs in language. Such reasons may be the clarity of the meaning dimension, contextual versatility, symmetry, word frequency, frequency of co-occurrence, pair-wise acquisition, and stylistic co-occurrence preferences.

Co-occurrence has been shown to be an important factor in semantically oriented antonym experiments such as elicitation experiments and judgement experiments (Paradis et al. 2009 and Willners and Paradis 2010). It is well known that some words tend to collocate by virtue of the fact that they are related in meaning, as is the case with antonyms, or because the words are members of more or less fixed expressions, e.g. nominal constructions (*horse doctor*, *apple pie*, *great white shark*) or idiomatic expressions (*sick and tired*, *safe and sound*). Likewise, some words tend to co-occur in certain contexts or genres (*social* and *political*). In this study the focus is on the question of the importance of co-occurrence frequency for antonyms as well as for unrelated word pairs. The purpose is to find out whether co-occurrence frequency in itself can produce a priming effect from one word (an adjective) to its antonym (another adjective).

The way we investigate this question is schematized in Table 1. In a visual lexical decision task, we look at priming effects in antonym pairs and unrelated adjective pairs with varying degrees of co-occurrence frequency in the British National Corpus (the BNC). The example pairs in Table 1 illustrate the basic design of the study. We assume that co-occurrence frequency and adjective frequency are independent of one another to a certain extent. The

* This work was supported by a generous grant from the Swedish Research Council (www.vr.se) and forms part of a project called *Contrast in language, thought and memory*. We are grateful to Simone Löhndorf for her help with the experiments and to Roy Liddle for his comments on an earlier draft of this article.

adjectives *near* and *distant* are themselves more frequent than *vertical* and *horizontal*. Nevertheless, as a pair of antonyms, *horizontal* and *vertical*, co-occur far more often (390 times) than *distant* and *near* (11 times). Similarly, the adjectives *little* and *nice* co-occur 544 times, whereas *busy* and *plain* co-occur only twice.

Table 1 here

Section 2 reviews existing studies on antonym canonicity. In Section 3, our focus is narrowed down to co-occurrence frequency of antonyms. Section 4 states the research objective and the aims of this study and is used as a transition into the present study which is presented in Section 5. The results are discussed in Section 6.

2. Antonymy and canonicity

Antonym canonicity has recently been the focus of attention in a number of different investigations, using both textual and experimental techniques (e.g. Paradis and Willners 2007, Jones 2007, Jones et al. 2007, Murphy et al. 2009, Paradis et al. 2009, Willners and Paradis 2010). Antonym canonicity or goodness of opposability is the extent to which antonyms are both semantically related and conventionalized as pairs in language (Murphy 2003: 31). We argue that 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 pairs (Paradis et al. 2009). Antonym canonicity concerns lexico-semantic pairings in language. The lexical aspect of canonicity concerns *which* words pairs are located *where* on an imagined scale from good to bad antonyms as measured in terms of participants' assessments and in terms of co-occurrence patterns in text, while the semantic side of the matter focuses on *why* some pairs might be considered better oppositions than others with reference to their conceptual set-up. In other words, while the lexical side of the coin concerns word forms, the semantic side of the matter concerns the characteristics of the meaning structure evoked by the word forms (Paradis 2010, Paradis and Willners 2011).

Paradis et al. (2009) and Willners and Paradis (2010) used English and Swedish adjectives respectively to measure which adjectives form part of strongly canonical antonymic relations and which adjectives have no strong candidate for this relationship. The method of data selection in both the English and the Swedish study was corpus-driven, sampled according to sentential co-occurrence frequencies. Two types of experiments were carried out using the pairings retrieved from the corpora – a judgement experiment and an

elicitation experiment.¹ In the first part of the study (the judgement experiment), participants were asked to rate the goodness of oppositeness of adjective pairs on a scale from 1 to 11. In the second part (the elicitation experiment), another group of participants were given one adjective for which they were asked to provide the best possible opposite. The hypothesis under investigation was 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 pairings

On the basis of the data set used in Paradis et al. (2009), it was shown that the adjectives that were deemed to be excellent antonym pairs by the participants in the judgement experiment were also the ones, among the words searched for, that were the most frequently occurring in the BNC both as individual words and as co-occurring pairs. A set of canonical antonyms that differed significantly from the rest in the judgement experiment was identified. This result was subsequently confirmed by the results of the elicitation experiment, which showed that canonical antonyms elicited significantly fewer opposites than other adjective pairs. More generally, a strong correlation between adjective frequency and the number of antonyms suggested by the participants was found (Spearman rho = -0.62, $p < .01$) (Paradis 2010). In other words, more frequent adjectives tend to elicit fewer different antonyms than less frequent adjectives. The actual frequency for these adjectives is a sign of the fact that they may qualify a large range of nominal meaning structures and are useful in a large range of contexts. The fact that they are very frequent also calls for more research on that parameter, hence the topic of this investigation. It was also shown in the judgement experiment that sequential order does not play any role in the participants' assessments of goodness of antonymy.

Another factor that seems to be of importance for the best pairings, judging from the experimental results, is the salience of the dimension. The dimension of which the canonical antonyms are representatives is salient if it is easily identifiable. For instance, the SPEED dimension underlying *slow-fast* is easily identifiable, while the dimension underlying say *rare-abundant*, *calm-disturbed*, *lean-fat*, *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 within monosemy do not prevent a word from participating in a canonical relation with another word, e.g. *light-dark* and *light-heavy*, and *narrow-wide* and *narrow-open*. Contextual versatility is a reflection of ontological versatility, i.e. 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.

For the sake of the investigation, two approaches to antonymy were set up as contrasting positions by Paradis et al. (2009): the *lexical categorical* approach and the *cognitive prototype continuum* approach. The former approach considers antonymy to be

¹ Both studies yielded the same general results. In this article we will restrict ourselves to reporting on the English data from Paradis et al. (2009).

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 are members. The model is context insensitive and static. Words either have antonyms or not. If they have antonyms they have one antonym. For instance, Miller and Fellbaum (1991: 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 is not the case according to Miller and Fellbaum (1991). The conceptual opposition in their model between, say, *ponderous* and *light* is mediated by *heavy*. Conceptual opposition is thus an effect of lexical relations rather than its cause. However, the experiments carried out by Paradis et al. (2009) and Willners and Paradis (2010) 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. This predicts 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 a conceptual basis: antonymy is a construal rather than a pre-stored representation. It is dependent on general cognitive processes such as comparison and profiling and relies on a binary configuration of a segment of content (Paradis et al. 2009, Paradis and Willners submitted, see also Cruse and Croft 2004 for a construal approach to antonymy in the Cognitive Linguistics framework and Murphy 2003 for a context-sensitive pragmatic approach). Adjectival meanings are evoked 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 the investigation carried out by Paradis et al. (2009) and Willners and Paradis (2010) indicate that strongly canonical pairings have lexical correlates, while the vast majority of antonyms have only associatively weak partners in situations where 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 and Andrew 1993). In the lexical categorical model different contexts do not affect the antonym, since the antonym of a word is not determined by context and sense, but is lexically driven. Finally, the

prototype continuum model is consistent with categorization in general (Taylor 2003).

In line with the reasoning and the findings of Murphy and Andrew (1993) and Murphy (2002), the theoretical implication of our investigation is that antonymy is primarily a conceptual relation, based on general knowledge-intensive cognitive processes. However, the investigations also point to the fact that a select group of antonyms are particularly strongly associated in memory. They are deemed to be superb examples of antonyms by participants in judgement experiments and there is strong agreement across participants in elicitation performance about the best antonym of a given word from this select group. For instance, even though it is easy to produce possible antonyms of *bad* (*satisfactory*, *beneficial*, *fine*, *obedient*), all of the experiment participants offered *good* as the best antonym of *bad* in the above-mentioned elicitation experiment. Pairings for which the participants suggested many different antonyms in the elicitation experiment are more likely to be contextually limited, i.e. not strongly routinized as pairs in our minds, or very weakly conventionalized, more generally, due to extreme genre or register restrictions.

In spite of the fact that the test items in both the judgement and the elicitation experiments were presented out of context, the experimental types relate more to the semantic side of the pairings than the formal side. This is also the case in other priming experiments such as Becker (1980). In his visual word recognition experiment, Becker used two types of stimuli: antonyms, such as *smart–dumb*, *dry–wet* and category-name/category-member pairs, such as *furniture–chair*, *dog–collie*. They were presented both as related, i.e. *smart–dumb*, and as unrelated, i.e. *smart–dry*, and the same design was used for the categorically related pairs. The cue-target materials were designed to produce a situation in which the participants were to predict consciously what a related target would be. The problem under investigation concerns the conditions under which a facilitating effect of an appropriate semantic context dominates and the conditions under which an interfering effect of an inappropriate context dominates. The experiment using antonyms produces a substantial facilitation effect and negligible interference, while the category-member relationship yielded only nominal facilitation but substantial interference. Like Paradis et al. (2009), Becker's investigation also shows that the order of the test items is of no importance. What is of importance, however, are the qualitative distinctions in terms of strength of relatedness. Both in the case of what Becker refers to as high-typicality antonyms (strongly canonical antonyms in our terminology) and high-typicality category-member pairings, relatively small interference effects but substantial facilitatory effects obtain. His conclusion is that the crucial factor is the type of stimuli. Facilitation dominance prevails for the condition that contains consistently strongly related test items, while interference dominance obtains when the test items are characterized by a wide range of semantic relationship strengths.

In sum, there are indications in the above-mentioned investigations that co-occurrence frequency is of importance for antonym canonicity. We take this as the springboard into the next section which concerns the current investigation of the role of frequency for strength of lexical affinity of antonym couplings.

3. Antonyms and frequency

There is widespread agreement in the literature that word frequency and semantic relatedness have facilitating effects in visual lexical decision performance (e.g. Becker 1979, 1980, Perea and Rosa 2002a, 2002b). Frequent words are recognized at higher speed than infrequent words, and targets preceded by related primes, e.g. *table-chair*, are known to be recognized faster than targets preceded by unrelated primes, e.g. *table-moon*. Similarly, as already mentioned above, canonical antonyms have been found to prime each other more strongly than other opposites or words that are related through category-membership, such as *furniture-chair*, *dog-collie* (Becker 1980). It has also been shown in semantic priming experiments that low-frequency targets produce larger priming effects than high-frequency targets (Becker 1979, Plaut and Booth 2000). In most of the literature, frequency refers to individual word frequencies rather than to co-occurrence frequency. There are a few exceptions to this in the priming literature, however. For instance, using different kinds of related words, such as synonyms and various types of category-membership relations, Spence and Owens (1990) demonstrate that co-occurrence frequency in text is significantly correlated with strength of association in priming experiments. In contrast to these findings, Estes and Jones (2009) show that co-occurrence frequency does not play a role in the explanation of integrative priming in expressions such as *lemon cake*, *horse doctor* and *plastic toy*.

The reasons for using a corpus-driven method for stimulus selection in Paradis et al. (2009) and Willners and Paradis (2010) and also in this study are twofold. One reason is to be able to establish the frequency levels for antonyms in naturally occurring, non-manipulated text and discourse. The other reason is to select test items for the experiments in a principled way using natural language since previous corpus studies have shown that textual evidence supports degrees of lexical canonicity. Charles and Miller (1989), Spence and Owens (1990), Justeson and Katz (1991, 1992), Fellbaum (1995) and Willners (2001) have established that members of pairs 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).

The same assumptions are made in corpus-based treatments of antonym co-occurrence in text (Jones 2002, 2007, Jones et al. 2007, Muehleisen and Isono 2009, Murphy et al. 2009). These studies concern the aspect of frequently co-occurring antonyms serving various contrasting discourse functions in text in constructional frames such as 'neither X nor Y', 'X instead of Y', the difference between X and Y'. The results attest to the fact that antonym pairs which are perceived to be good opposites occur frequently in such discursive frames. Similar studies have been performed on data from the CHILDES database (Jones and Murphy 2005, Murphy and Jones 2008, Tribushinina (submitted). For instance, using the American English component of the database, Murphy and Jones (2008) observe that children use antonyms at earlier ages than experimental studies have shown and they also use

antonyms for mostly the same discursive purposes as adult speakers do. It has also been shown that lexical access and various levels of lexical priming may influence word choice and word prediction (Yap et al. 2009) and lexical priming may also be a triggering factor for speech errors (Söderpalm 1979, Gainotti et al. 1983, Varley 1991, Samson et al. 2007).

As was previously mentioned, co-occurrence has also been used in integrative priming experiments, i.e. testing the relations between two nouns, where the first noun is a modifier of the second noun and thereby designates a subcategory of the head noun in that they jointly name the category (Estes and Jones 2009). For instance, the concepts for *table* and *vase* may be integrated through a relation of location (*table vase*), through causation (*rope burn*), through composition (*copper pot*), through time (*winter holiday*), through function (*sketch pen*), through partonymy (*bear paw*), through topic (*cowboy film*) and through production (*wind power*). Estes and Jones (2009) examined McKoon and Ratcliff's (1992) argument that the frequency of co-occurrence in samples of written text is the best estimate of prime-target familiarity but found that co-occurrence in text did not explain integrative priming. Instead, Estes and Jones (2009) explain the mechanism of integrative priming using role assignment, e.g. location or causation, and propose that role typicality, relation plausibility and compound familiarity are crucial factors.

4. Research objective and aims

The results of the studies reviewed in the previous sections suggest a correlation between co-occurrence frequency and 'goodness of antonymy' of the form-meaning pairings. In line with these results, we assume that antonyms that co-occur often tend to be assessed as better pairs than antonyms that do not occur often. Even though more frequent adjectives generally co-occur more often than less frequent adjectives, we make the assumption that both co-occurrence frequency and adjective frequency may have independent facilitating effects on word recognition.

Our aim in this study is therefore to find priming effects that are due to co-occurrence frequency independently of individual word frequency. We are hereby able to evaluate the relative importance of co-occurrence frequency for the antonymic relations on the hypothesis that frequently co-occurring antonyms are more strongly conventionalized as pairings in text and discourse. The prediction that follows from the hypothesis is that frequency of co-occurrence of pairings speeds up word recognition. Indirectly, the outcome of such an experiment allows us to evaluate the two approaches to antonymy, the *lexical categorical* approach and the *cognitive prototype continuum* approach that were described in Section 2. The second aim is to establish whether co-occurrence frequency can cause a priming effect when two words are semantically unrelated. For instance, the adjective *nice* will be recognized faster when it is preceded by *little*, since these two words often co-occur, whereas there is no facilitating effect for the word *plain* when it is preceded by *busy*, since these two adjectives are unrelated (see Table 1). The inclusion of semantically unrelated word

pairs also allows us to examine possible interaction effects between co-occurrence frequency and semantic relatedness. At a more general level, co-occurrence frequency has rarely been used as a factor in priming experiments. Our results, therefore, complement existing findings on priming, and should therefore have implications for models of priming.

5. The experiment

This investigation combined corpus data with behavioural data. We carried out a visual lexical decision experiment for which we selected material, based on frequency information obtained from the BNC. In a visual lexical decision task, words and nonsense words were presented to participants on a computer screen. The participants’ task was to press one of two buttons on a button box for a real word and the other button for a nonsense word. They were encouraged to respond quickly and accurately. Their reaction times were recorded, as well as their errors (incorrect button presses). Notable factors that influence reaction times are, among others, word frequency (high-frequency words are recognized faster than low-frequency words) and word length (short words are recognized faster than long words (cf. New et al. 2005)). We controlled for these factors in order to avoid interferences with potential co-occurrence frequency.

A total of 20 participants were tested: 17 women and 3 men, aged 19 to 40 years (average age was 24.75 years). Most of the participants were exchange students at the universities of Lund and Växjö in Sweden. They were recruited through the International Offices at the two universities. All of them had English as their native language from Australia (1), Canada (1), South Africa (1), Great Britain (7) and the United States of America (10). The reason for choosing the BNC as a source for data retrieval was a matter of the size of the corpus rather than the fact that the data are British English. We expect educated native speakers from any English-speaking part of the world to be equally familiar with the test items we used for the experiment, which are all frequent, ‘common core’ adjectives in all of the above dialects (see Appendix A). We obtained the participants’ written consent to use their input for our investigation. At the end of the experiment, they received a lottery scratch card as compensation for their time and effort.

Table 2

Materials

The adjective pairs that we used as primes and targets in the current experiment were selected on the basis of frequency counts in the BNC. The adjectives that were considered to be suitable items were all relatively common, with a frequency of at least 1,000 occurrences in the BNC. We compiled a list of all the adjective pairs that co-occurred within a sentence. The list consisted of 422,499 pairs with

token frequencies ranging from 1 to 3,946 (a total token frequency of 4,454,280). Table 2 shows the ten most frequent adjective pairs. From the list, we manually identified 233 antonym pairs of which we chose 60 for the experiment, and decided which of the two adjectives would serve as the prime and which would be the target. Subsequently, we chose another 60 unrelated adjective pairs from the same list and matched these with the antonym pairs so that the targets had approximately equal length and frequency (see Table 3). No adjective occurred more than once in the experiment. The adjective pairs are shown in Figure 1. Note that the antonym pairs appear to have slightly higher co-occurrence frequencies than the unrelated. We will return to this issue in the analysis.

Figure 1

To the 240 adjectives, we added 287 phonotactically appropriate but non-existing English words, e.g. *goast*, *solt*, *voddle* and *foose*, yielding a total of 527 items. The item pairs were ordered randomly and mixed with the nonsense words in such a way that there were always one or two nonsense words between a target and a prime, and either zero, one, or two nonsense words between a prime and a target. The gap between primes and targets was varied randomly in order to make the appearance of the target word unpredictable for the participants.

Experimental procedure

The presentation of the stimuli and the collection of the participant responses were controlled using E-prime software (Schneider, Eschman and Zuccolotto 2001). An experimental trial consisted of a time interval of 1500 ms during which the word was presented on the computer screen followed by a 500 ms empty screen. The participants were instructed to press either one of two buttons on a button box for a real word or a nonsense word. Only responses during the time that the item was on the screen were recorded. Responses given outside that interval were counted as errors.

The experiment was divided into five blocks. The first block consisted of 10 practice items (not included in the 527 experimental items) to make sure that the participants understood and followed the instructions correctly. After that, the experimental items were presented in four approximately equally long blocks, giving the participants the opportunity to relax for a short while, and then continue with the next block when they felt ready to do so. The participants were tested in a silent room.

Analysis

In the analysis, we focused exclusively on the reaction times to the target adjectives. We were mainly interested in two factors: co-occurrence frequency (a continuous variable) and relatedness (antonyms or unrelated). In addition to these two main factors, we also looked at potential confounding factors: the individual frequency of the target, the length of the target, (the number of letters), the trial number (the moment at which a particular item appeared in the

experiment) and the gap size between prime and target (0, 1 or 2 nonsense words). We had two concerns that were importance for our analysis of the data. One was the effect of possible interference of the confounding factors and the main predictors. Another concern was that co-occurrence frequency was slightly unbalanced across the experimental items, with the antonyms having a higher average co-occurrence frequency than the unrelated adjectives (see Table 3).

Table 3

In order to overcome these concerns, we disentangled the effect of the main predictors and the effects of the confounding factors and analyzed the results by fitting a multilevel model to the data. Due to its flexibility, the multilevel approach had several advantages. First, the effects of the main predictors were tested while controlling for the effects of the confounding predictors. Second, unbalanced data (unless severely unbalanced) were unproblematic for the analysis (Singer and Willett 2003). Finally, frequency and word length were continuous variables. This would have been problematic for an analysis of variance but was not for the multilevel approach where predictors may be categorical or continuous.

In the model, reaction time was the dependent variable, which is predicted from the main factors, the confounding factors, and two random factors (i.e., participant and stimulus word). We measured the interaction of co-occurrence frequency and relatedness but also looked at their main effects. The predictions were straightforward: we expected a (semantic) priming effect for antonyms but not for unrelated adjective pairs.

Results

The first row in Table 4 shows the percentages of correct responses for the targets. Overall, the participants made remarkably few errors. The overall percentages of correct responses approached 99% both for the antonymic and the unrelated targets. Incorrect responses were excluded from further analysis. The second row of the table shows the average reaction times for the antonymic and the unrelated targets. In line with what we expected, responses to antonymic targets were faster than those for the unrelated targets by nearly 20ms.

Table 4

The relation between co-occurrence frequency and reaction times is shown in Figure 2. For this figure and for the subsequent analysis, the reaction times and co-occurrence frequency were log-transformed in order to reduce skewness and the effect of outliers (Baayen 2008: 71). The slopes of the regression lines in the figure suggest a very slight negative relationship for the antonym pairs, and a very slight positive relationship for the unrelated adjective pairs.

Figure 2

The analysis started with an initial model containing all predictors: trial, item frequency, item length, lag size (0, 1 or 2 words), and the joint effects (i.e., the interaction plus the main effects) for relatedness and co-occurrence frequency. The results are given in Table 5. This table gives the estimates of the effects, together with their standard errors. The sign of a coefficient (plus or minus) indicates the direction of the relationship between the predictor and the dependent variable (positive or negative). The last column shows the accompanying *t*-values, indicating whether the effect of the predictor is significant or not. Since the number of observations is rather large (over 2,000), the *t*-values may be interpreted as *z*-scores, and thus absolute values larger than 1.96 may be considered significant with $p < .05$ (cf. Baayen 2008: 248).

Table 5

The *t*-values given in the rightmost column suggest significant effects of trial, target length, and target frequency, but not of relatedness, lag size, co-occurrence frequency nor of their interaction. In the next step, we excluded the interaction term and tested the individual effects of relatedness and co-occurrence frequency by fitting two new models excluding either of these two factors. The deviance statistics show that the omission of co-occurrence frequency as a predictor does not give a significantly worse fit ($X^2 = 3.609$, $df = 2$, $p = 0.165$), but the omission of the relatedness predictor does ($X^2 = 8.148$, $df = 2$, $p = 0.017$).

We then continued fitting four simpler models, each without the co-occurrence predictor and without any of the four remaining predictors. The results show that lag size can be omitted from the model without leading to a significantly worse fit ($X^2 = 0.005$, $df = 1$, $p = 0.943$), but none of the other predictors (target frequency: $X^2 = 3.988$, $df = 1$, $p = 0.046$; target length: $X^2 = 14.007$, $df = 1$, $p = 0.000$; trial: $X^2 = 5.980$, $df = 1$, $p = 0.014$) can be omitted. The coefficients of this final model with the four predictors are displayed in Table 6.

There is a negative relationship between trial and reaction times, indicating that the participants tend to become slower towards the end of the experiment. Similarly, there is a negative relationship between reaction times and the frequency of the target, indicating that the participants respond to high-frequency targets faster than to

low-frequency targets. The positive coefficients for relatedness and target length indicate that responses to antonyms are faster than those to unrelated adjectives, and that responses to longer targets are slower than responses to shorter targets.

Table 6

6. Discussion

As stated in the introduction, previous research using experimental methods as well as corpus techniques show that the relation of antonyms has special status in terms of semantic priming effects and in terms of co-occurrence in text and discourse and that some very frequent and frequently co-occurring adjectives stand out as excellent members of the category. Antonyms prime each other more often than other related words and they co-occur in sentences significantly more often than other words. It has also been shown that the most strongly co-occurring antonyms in text are also the ones that people judge to be excellent antonyms in judgement experiments. In elicitation experiments these adjectives elicit only one or a couple of antonyms, which we interpret as an indication of ‘goodness of antonymy’ both in terms of semantic relatedness and lexical association. On the basis of their priming experiments, Spence and Owens point out that frequency of co-occurrence is a function of association strength, and unlike the data used in Charles and Miller (1989), Justeson and Katz (1991, 1992), Paradis et al. (2009) and Willners and Paradis (2010), their co-occurrence data extend beyond the borders of the sentence. They report that priming effects do not seem to decline until some 200 words beyond the stimulus word. However, there are also studies with results that point in a different direction. Estes and Jones (2009) investigated whether frequency of co-occurrence in large samples of written language provides better estimates of prime-target familiarity. They, too, used both adjacent words and global co-occurrence in text, using Latent Semantic Association (LSA) for the latter type of co-occurrence. They found no crucial priming effects related to co-occurrence frequency, neither local nor global.

In spite of the difference of semantic relatedness across the above studies and the different types of experiments using co-occurrence frequency, the contradictory results of Estes and Jones (2009) led us to look more closely into whether goodness of antonymy is a matter of strength of relatedness or lexical association and to what extent frequency of co-occurrence plays a role for the status of strongly canonical antonyms. In contrast to our own, more semantically oriented previous experiments, we designed this experiment as a lexical recognition task in order to put the spotlight on the lexical side of the matter. We argued that if it is the case that antonyms which co-occur often in text and discourse are considered better antonyms than antonyms that do not co-occur often, then we would expect such form-meaning pairings to be more strongly

entrenched in memory than less frequently co-occurring pairs of form-meaning pairings. As a result, the presentation of one member of a frequently co-occurring pair will facilitate the recognition of the other member. This will not be the case if the members of a pair do not co-occur often.

In order to investigate this state of things, we selected pairs of adjectives with varying degrees of co-occurrence frequencies for the BNC. Half of these pairs were antonyms, and the other half were not related in meaning. We carried out a visual decision experiment to see whether we would find evidence of priming from one member of a pair to the other. The results showed that the recognition of antonymic targets was indeed facilitated by their primes, which was not the case for the unrelated pairings. We also found a facilitatory effect of target word frequency on reaction times: frequent targets were recognized faster than less frequent targets. Crucially, however, we did not find that frequently co-occurring antonyms, such as *horizontal* and *vertical*, primed each other more than less frequently co-occurring antonyms, such as *distant* and *near*. This means that when participants see *horizontal*, they lower the threshold for *vertical*. This effect is not due to the fact that *horizontal* and *vertical* co-occur frequently but because they are semantically related. The facilitation of the target word also happens in the case of less frequently co-occurring antonyms such as *distant* and *near*. Facilitation is thus not likely to be a consequence of lexical association, and it cannot be attributed to familiarity either, but has a semantic basis. Similarly, in the case of the unrelated test items, we found no priming effect and no facilitating effect due to frequency of co-occurrence either.

In our experiment we manipulated lag size between the prime and the target, which is a different procedure from Estes and Jones' (2009) experiment on co-occurrence and compound nouns. They used Stimulus Onset Asynchronies (SOA) to investigate different types of priming and found that both associative priming and semantic priming are observed at short (i.e. <300ms) and intermediate SOAs (approximately 300ms-800ms). Associative priming continues to increase in magnitude across longer SOAs (i.e. $\geq 1,000$ ms), whereas semantic priming tends to dissipate at those later SOAs (e.g. Perea and Rosa 2002). Instead of using SOAs, we manipulated the number of words in between stimuli. The stimuli were presented with either zero, one or two nonsense words between the prime and the target. The reason why we varied the distance was mainly to avoid testing fatigue and monotony and to prevent the participants from developing experiment strategies.

The analysis started with an initial model containing all predictors, item frequency, item length, lag size (0, 1 or 2 words), and the joint effects (i.e., the interaction plus the main effects) for relatedness and co-occurrence frequency. The outcome of the analysis suggests significant effects of trial, target length, and target frequency, but not of relatedness, lag size, co-occurrence frequency nor of their interaction. It would be interesting to repeat our experiment using varying SOAs instead of intervening items.

Other potentially confounding variables may be participants' level of word knowledge. Yap et al. (2009) isolate various variables in order to be able to show a more nuanced picture and in order to better explain the relation between priming effects

and word frequency. Through a multi-stage model of lexical processing consisting of a lexical (perceptual, orthographic in this case) module and a semantic (cognitive) module, they measure processing differences across participants on the basis of their level of vocabulary knowledge. They show that the joint effects of semantic priming and word frequency are critically dependent on the participants' level of vocabulary knowledge. Yap et al. also show that semantic priming and word frequency do not always interact. Participants with less vocabulary knowledge show larger priming effects than participants with higher vocabulary knowledge. The priming effects among the former group of participants are particularly pronounced for low-frequency targets. They argue that the result is consistent with the idea of a flexible lexical processing system in which participants' performance is optimized by relevant contextual information. In contrast, the lexical processing system of participants with higher vocabulary knowledge appears to be more modular in nature, i.e. the prime provides a head-start that is independent of how difficult the target is. We do not think that participants' vocabulary knowledge plays a role in our data, because the participants are all well-educated speakers with university education and all the test items are common words in English, occurring at a rate of more than 1,000 times in the corpus.

Priming is typically attributed to strength of lexical association and semantic relatedness. Frequency effects are obtained: frequent words are recognized faster than rare words. Similarly, if words preceded by semantically related primes are recognized faster than unrelated primes, we have a semantic effect. This distinguishes between association priming, which is caused by lexical associative strength, and semantic priming, which is due to semantic similarity. There are a number of theoretical models available for the explanation of lexical priming in the literature: association models such as spreading activation (e.g. Anderson 1983), and expectancy models (e.g. Becker 1980). There are also more semantically oriented models such as the distributed representation model (e.g. Plaut and Booth 2000), which assumes priming to occur when words have overlapping patterns of activation of semantic features represented in different parts of the brain. Moreover, there is the compound cue model (Ratcliff and McKoon 1988), which posits that a prime forms a cue which is matched against long-term memory, in which case the prime-target relation is a result of the extent to which they are associated in memory, or in Ratcliff and McKoon's terminology, the extent of the familiarity of the pairing. In their model, co-occurrence in text has the status of being the best predictor of strength of familiarity (McKoon and Ratcliff 1992).

This view has been challenged by Estes and Jones (2009) who showed that lexical priming also occurs among unassociated, dissimilar and unfamiliar concepts such as *horse* and *doctor*. Such priming is said to occur when a prime word can be easily integrated with the meaning of a target word to create a unitary representation. It was also shown that integrative priming was different from associative and semantic priming but comparable to them in terms of prevalence across the participants as well as magnitude within participants. Estes and Jones (2009) argue that this finding constitutes a challenge to models such as spreading activation, distributed representation, expectancy, episodic retrieval and

compound cue models, and they suggest that it can be explained by a role activation model of relational integration. In spite of the fact that the test items in Estes and Jones (2009) were compound concepts, unlike our semantically related antonyms, the result of that experiment is similar to ours in that they did not obtain any co-occurrence frequency effects.

7. Conclusion

This study has two important results. It confirms the hypothesis and previous findings that antonymic targets are facilitated by their primes, but it does not confirm the hypothesis that frequency of co-occurrence facilitates word recognition, either for antonyms or for unrelated adjectives. This means that there is a relatedness effect but no co-occurrence frequency effect, which in turn means that priming cannot be attributed to lexical association. The prime-target effect we obtain is a semantic effect, indicating that conceptual opposition is the *cause* of lexical relation rather than the other way round the *effect* of the lexical relation. This piece of evidence lends support to a conceptual rather than a lexical approach to antonymy.

Table 1. The experimental conditions

| Frequency of co-occurrence | Antonyms | Unrelated word pairs |
|----------------------------|----------------------------|----------------------|
| High | <i>horizontal-vertical</i> | <i>little-nice</i> |
| Low | <i>near-distant</i> | <i>busy-plain</i> |

Table 2: Most frequent adjective pairs in the BNC.

| | | |
|-----------|-----------|-------|
| new | old | 3,946 |
| economic | social | 3,765 |
| black | white | 3,498 |
| new | other | 3,493 |
| economic | political | 2,816 |
| political | social | 2,787 |
| other | social | 2,503 |
| large | small | 2,453 |
| different | other | 2,377 |
| local | other | 2,259 |

Table 3: Descriptive statistics for the target words. Frequencies are log-transformed values.

| | Antonyms | Unrelated |
|-------------------------|----------|-----------|
| co-occurrence frequency | 4.97 | 3.43 |
| target frequency | 8.77 | 8.99 |
| target length (letters) | 5.97 | 6.02 |

Table 4: Percentages correct and average Reaction Times.

| | Antonyms | Unrelated |
|-------------|----------|-----------|
| Correct (%) | 98.6 | 98.5 |
| RT (ms) | 572 | 591 |

Table 5: Initial[e1] model.

| | coefficient estimate | standard error | <i>t</i> |
|---------------|----------------------|----------------|----------|
| trial | -0.0002 | 0.0001 | -2.45* |
| relatedness | 0.0058 | 0.0376 | 0.15 |
| co-occurrence | 0.0030 | 0.0063 | 0.48 |

| | | | |
|-------------------------------------|---------|--------|------------|
| frequency (log) | | | |
| co-occurrence freq × relatedness | 0.0082 | 0.0082 | 1.01 |
| target frequency (log) | −0.0208 | 0.0079 | − 2.64* |
| lag size | −0.0016 | 0.0079 | −0.20 |
| target length | 0.0114 | 0.0033 | 3.44* |

Table 6: final
model

| | coefficient estimate | standard error | <i>t</i> |
|---------------------------|-------------------------|-------------------|------------|
| trial | −0.0002 | 0.0001 | − 2.52* |
| relatedness | 0.0270 | 0.0127 | 2.13* |
| target frequency (log) | −0.0134 | 0.0068 | − 1.98* |
| target length | 0.0124 | 0.0033 | 3.80* |

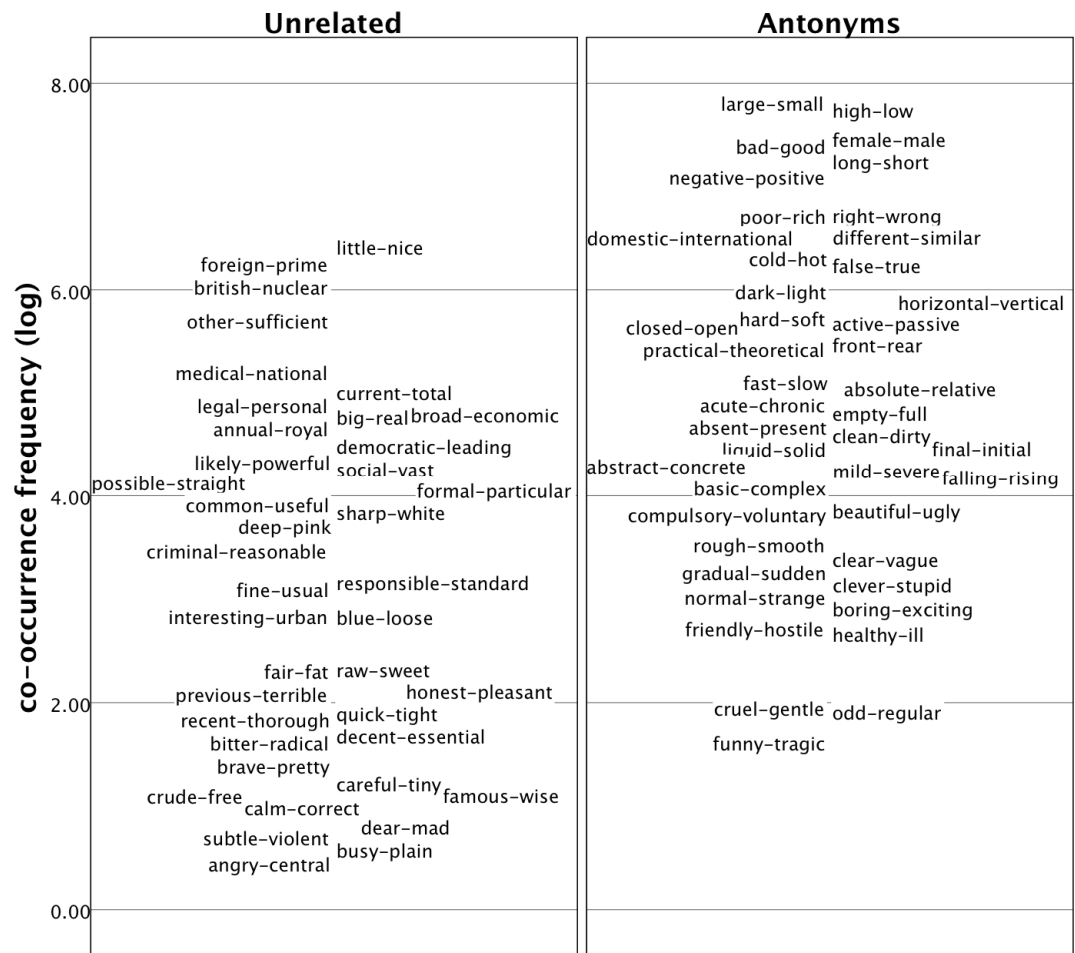


Figure 1: Experimental item pairs. The vertical positions of the adjective pairs correspond to their approximate frequency of co-occurrence.

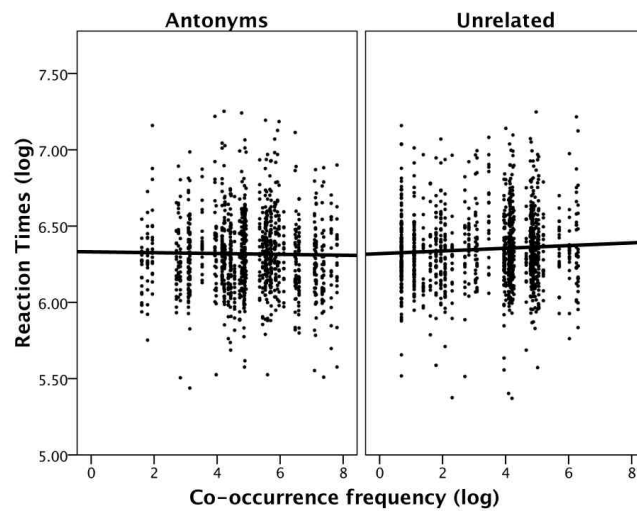


Figure 2: Relationship between co-occurrence frequency and reaction times.

| Appendix A: Experimental items with co-occurrence frequencies. | | | |
|---|------|----------------------|------|
| <i>Antonyms</i> | | <i>Unrelated</i> | |
| large-small | 7.81 | little-nice | 6.30 |
| high-low | 7.62 | foreign-prime | 6.25 |
| bad-good | 7.39 | british-nuclear | 6.03 |
| female-male | 7.33 | other-sufficient | 5.70 |
| old-young | 7.15 | medical-national | 5.20 |
| long-short | 7.12 | great-private | 5.05 |
| negative-positive | 7.09 | appropriate-relevant | 5.04 |
| right-wrong | 6.60 | effective-simple | 5.02 |
| poor-rich | 6.59 | big-real | 4.97 |
| different-similar | 6.51 | early-popular | 4.93 |
| cold-hot | 6.48 | current-total | 4.89 |
| false-true | 6.12 | legal-personal | 4.88 |
| horizontal-vertical | 5.97 | financial-serious | 4.88 |
| strong-weak | 5.92 | general-late | 4.86 |
| dark-light | 5.86 | major-rural | 4.84 |
| domestic-international | 5.83 | necessary-proper | 4.83 |
| hard-soft | 5.72 | immediate-political | 4.81 |
| difficult-easy | 5.72 | natural-public | 4.80 |
| spoken-written | 5.70 | broad-economic | 4.79 |
| permanent-temporary | 5.68 | annual-royal | 4.65 |
| dry-wet | 5.60 | heavy-main | 4.26 |
| closed-open | 5.57 | democratic-leading | 4.23 |
| active-passive | 5.56 | deaf-special | 4.22 |
| practical-theoretical | 5.54 | fresh-whole | 4.20 |
| alive-dead | 5.53 | likely-powerful | 4.17 |
| new-used | 5.45 | available-safe | 4.16 |
| front-rear | 5.35 | social-vast | 4.14 |
| empty-full | 4.92 | black-flat | 4.14 |
| acute-chronic | 4.88 | formal-particular | 4.13 |
| absolute-relative | 4.88 | familiar-important | 4.11 |
| multiple-single | 4.88 | complete-human | 4.09 |
| narrow-wide | 4.87 | possible-straight | 4.01 |
| cheap-expensive | 4.86 | common-useful | 3.97 |
| fast-slow | 4.85 | deep-pink | 3.95 |
| global-local | 4.80 | sharp-white | 3.95 |
| classical-modern | 4.78 | criminal-reasonable | 3.47 |
| thick-thin | 4.72 | fine-usual | 3.09 |
| absent-present | 4.55 | responsible-standard | 3.04 |
| liquid-solid | 4.44 | interesting-urban | 2.83 |
| clean-dirty | 4.43 | blue-loose | 2.71 |
| final-initial | 4.42 | fair-fat | 2.30 |
| happy-sad | 4.36 | previous-terrible | 2.08 |
| mild-severe | 4.25 | raw-sweet | 2.08 |
| basic-complex | 4.23 | honest-pleasant | 1.95 |
| abstract-concrete | 4.22 | recent-thorough | 1.95 |
| falling-rising | 4.16 | quick-tight | 1.79 |
| cool-warm | 4.16 | decent-essential | 1.79 |
| beautiful-ugly | 3.97 | bitter-radical | 1.61 |
| compulsory-voluntary | 3.93 | brave-pretty | 1.39 |

| | | | |
|------------------|------|----------------|------|
| rough-smooth | 3.53 | calm-correct | 1.10 |
| normal-strange | 3.14 | crude-free | 1.10 |
| gradual-sudden | 3.14 | careful-tiny | 1.10 |
| clever-stupid | 3.09 | famous-wise | 1.10 |
| clear-vague | 3.09 | dear-mad | 0.69 |
| boring-exciting | 2.83 | angry-central | 0.69 |
| healthy-ill | 2.77 | subtle-violent | 0.69 |
| friendly-hostile | 2.71 | busy-plain | 0.69 |
| cruel-gentle | 1.95 | guilty-proud | 0.69 |
| odd-regular | 1.79 | causal-genuine | 0.69 |
| funny-tragic | 1.61 | blind-silly | 0.69 |

References

- Anderson, John R. 1983. A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior* 22. 261-295.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
- Becker, Curtis. A. 1979. Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 5. 252-259.
- Becker, Curtis. A. 1980. Semantic context effects in visual word recognition: an analysis of semantic strategies. *Memory and Cognition* 8. 493-512.
- Divjak, Dagmar and Stefan Th. Gries. 2008. Clusters in the mind? Converging evidence from near synonymy in Russian. *The Mental Lexicon* 3(2). 188-213.
- Charles, Walter G. and George A. Miller. 1989. Contexts of antonymous adjectives. *Applied Psycholinguistics* 10(3). 357-375.
- Estes, Zachary and Lara Jones. 2009. Integrative priming occurs rapidly and uncontrollably during lexical processing. *Journal of Experimental Psychology* 138(1). 112-130.
- Fellbaum, Christiane. 1995. Co-occurrence and antonymy. *International Journal of Lexicography* 8(4). 281-303.
- Gainotti, Guido, Gabriele Miceli, M. Caterina Silveri and Giampiero Villa. 1983. The production of morphological and lexical opposites in aphasia. *Neuropsychologia* 21(6). 693-697.
- Herrmann, Douglas J., Roger Chaffin, Margaret P. Daniel and R. S. Wool. 1986. The role of elements of relation definition in antonym and synonym comprehension. *Zeitschrift für Psychologie* 194. 133-153.
- Jones, Steven. 2002. *Antonymy: a corpus-based perspective*. London: Routledge.
- Jones, Steven. 2007. 'Opposites' in discourse: a comparison of antonym use across four domains. *Journal of Pragmatics* 39(6). 1105-1119.
- Jones, Steven and M. Lynne Murphy. 2005. Using corpora to investigate antonym acquisition. *International Journal of Corpus Linguistics* 10(3). 401-422.
- Jones, Steven, Carita Paradis, M. Lynne Murphy and Caroline Willners. 2007. Googling for 'opposites': a web-based study of antonym canonicity. *Corpora* 2(2). 129-154.
- Justeson, John S., and Katz, Slava M. 1991. Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistics* 17. 1-19.

- Justeson, John S., and Katz, Slava M. 1992. Redefining antonymy: the textual structure of a semantic relation. *Literary and Linguistic Computing* 7. 176-184.
- McKoon, Gail and Ratcliff, Roger. 1992. Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18. 1155-1172.
- Miller, George A. and Christiane Fellbaum. 1991. Semantic networks of English. *Cognition*, 41, 197-229.
- Muehleisen, Victoria and Maho Isono. 2009. Antonymous adjectives in Japanese discourse. *Journal of Pragmatics* 41(11). 2185-2203.
- Murphy, Gregory L. and Jane M. Andrew. 1993. The conceptual basis of antonymy and synonymy in adjectives. *Journal of Memory and Language* 32. 301-319.
- Murphy, Gregory L. 2002. *The big book of concepts*. The MIT Press: Cambridge (MA).
- Murphy, M. Lynne. 2003. *Semantic relations and the lexicon*. Cambridge: Cambridge University Press.
- Murphy, M. Lynne and Steven Jones. 2008. Antonymy in children's and child directed speech. *First Language*, 28(4). 403-30.
- Murphy, M. Lynne, Carita Paradis, Caroline Willners and Steven Jones. 2009. Discourse functions of antonymy: a cross-linguistic investigation. *Journal of Pragmatics*, 41(11). 2159-2184.
- New, Boris, Ludovic Ferrand, Christophe Pallier and Marc Brysbaert. 2005. Re-examining the word length effect in visual word recognition: new evidence from the English lexicon project. *Psychonomic Bulletin and Review* 13. 45-52.
- Paradis, Carita and Caroline Willners. 2007. Antonyms in dictionary entries: methodological aspects. *Studia Linguistica* 61(3). 261-277.
- Paradis, Carita and Caroline Willners. 2011. Antonymy: from convention to meaning-making. *Review of Cognitive Linguistics* 9(2).
- Paradis, Carita, Caroline Willners and Steven Jones. 2009. Good and bad opposites: using textual and experimental techniques to measure antonym canonicity. *The Mental Lexicon* 4(3). 380-429.
- Paradis, Carita. 2010. Good, better and superb antonyms: a conceptual construal approach. *The Annual of Texts by Foreign Guest Professors*. Charles University in Prague, Faculty of Philosophy and Arts.
- Plaut, David C. and James R. Booth. 2000. Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review* 107. 786-823.
- Perea, Manuel and Eva Rosa. 2002a. The effect of associative and semantic priming in the lexical decision task. *Psychological Research* 66. 180-194.
- Perea, Manuel and Eva Rosa. 2002b. Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asynchronies? *Acta Psychologica* 110. 103-124.
- Samson, Dana, Catherine Connolly and Glyn W. Humphreys. 2007. When "happy" means "sad": neuropsychological evidence for the right prefrontal cortex contribution to semantic processing. *Neuropsychologia* 45(5). 896-904.
- Schneider, Walter, Amy Eschman and Anthony Zuccolotto. 2001. *E-prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.

- Spence, Donald P. and Kimberly C. Owens. 1990. Lexical co-occurrence and association strength. *Journal of Psycholinguistic Research* 19. 317-330.
- Singer, Judith and John Willett. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press.
- Söderpalm, Ewa. 1979. Speech errors in normal and pathological speech. *Travaux de l'Institut de Linguistique de Lund* 14.
- Taylor, John R. 2003. *Linguistic Categorization*. Oxford: Oxford University Press.
- Tribushinina, Elena. (submitted). Spatial adjectives in Dutch child language: towards a usage-based model of adjective acquisition. In Carita Paradis, Jean Hudson, and Ulf Magnusson (eds.), *Conceptual Spaces and the Construal of Spatial Meaning. Empirical evidence from human communication*. Under consideration for publication by Oxford: Oxford University Press.
- Varley, Rosemary. 1991. Reference, sense and antonymy in the assessment of lexical semantic abilities in aphasia. *Aphasiology* 5(2). 149-170.
- Willners, Caroline. 2001. Antonyms in Context. A Corpus-based Semantic Analysis of Swedish Descriptive Adjectives. *Travaux de l'Institut de Linguistique de Lund* 40.
- Willners, Caroline and Carita Paradis. 2010. Swedish opposites: a multi-method approach to goodness of antonymy. In Petra Storjohann (ed.), *Semantic relations: Theoretical and practical perspectives*, 15-48. Amsterdam: John Benjamins.
- Yap, Melvin, Chi-Shing Tse and David Balota. 2009. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* 61. 303-325.

|