Modelling Functional Priming and the Associative Boost

Scott McDonald (scottm@cogsci.ed.ac.uk)
Will Lowe (will@cogsci.ed.ac.uk)
Centre for Cognitive Science
University of Edinburgh, 2 Buccleuch Place
Edinburgh, EH8 9LW Scotland

Abstract

Using an auditory semantic priming paradigm, Moss, Ostrin, Tyler and Marslen-Wilson (1995, Experiment 2) demonstrated facilitation for category coordinates and functionally-related stimuli both with and without the additive effect of normative association strength. In this paper we replicate these results computationally using a corpus-derived Contextual Similarity measure. In Experiment 1 we consider the adequacy of the Contextual Similarity measure in accounting for Moss et al.'s results, and discuss how functional and categorical semantic relations are represented in corpus-based approaches to lexical semantics. We also offer an explanation for how the Contextual Similarity measure succeeds in replicating the additive effect of association strength on semantic priming without postulating a qualitatively different mechanism for associative priming. We then investigate why previous corpus-based approaches (Lund, Burgess & Atchley, 1995) have failed to produce similar results. We argue that this is because vector representations partly encode temporal co-occurrence information. This explanation is tested in Experiment 2.

Introduction

To understand the processes involved in human language comprehension it is important to determine what kind of information is available to the word recognition system. The widely-used semantic priming paradigm provides a minimal linguistic context, typically a single word, for investigating the factors that influence lexical processing (Neely, 1991). Although more than 25 years of priming research has shown that the presence of a related prime word tends to speed lexical decision to a target word, the type of relation between the prime and target words necessary to produce the effect is still under dispute.

Lexical Relations that Support Priming

The vast majority of semantic priming studies have concentrated on investigating words in a *taxonomic* relation. Materials typically consist of category coordinates, such as ⟨cat, dog⟩, though Neely (1991) reviews work examining superordinate and subordinate category relations between prime and target. Moss, Ostrin, Tyler and Marslen-Wilson (1995) point out that functionally related words, where the referents are related in ways that can be described in non-taxonomic terms, such as the instrument pair ⟨hammer, nail⟩ or the script relation ⟨restaurant, wine⟩, have often been assigned to the normatively associated[1] but *non*-semantic condition in the experimental design, thus confounding semantic and associative relations.

Priming has been observed between word pairs that are semantically related only (e.g. ⟨dance, skate⟩; Fischler, 1977), semantically related and normatively associated (e.g. ⟨dish, plate⟩; Moss et al., 1995), or merely frequently co-occur in text (e.g. ⟨hospital, baby⟩; McKoon & Ratcliff, 1992). The presence of priming for semantically related words in the absence of an associative relation has been particularly controversial (McRae & Boisvert, 1996; Shelton & Martin, 1992).

Moss et al. address these issues in three priming experiments by separately manipulating three factors: normative association (associated, non-associated), type of semantic relationship (category coordinate, functional), and semantic relatedness (related, unrelated). In the context of auditory presentation at least, a priming effect was observed for both category coordinates and functionally-related items. They conclude that functional information *is* accessed during word recognition. Furthermore, there was a priming effect both with *and* without the presence of normative association. By showing that priming occurred for functionally related but non-associated word pairs, Moss et al. uncovered a new source of information affecting the word recognition system. The second important finding from this set of experiments was the interaction between the association and relatedness factors: the presence of normative association resulted in a significantly larger priming effect. This additive effect was called the 'associative boost' (Moss et al., 1995).

Co-occurrence Statistics and Lexical Semantics

Recently there has been considerable interest in the psychological modelling of lexical semantic phenomena using computational techniques from statistical natural language processing (eg. Landauer & Dumais, 1997).

In this approach, the semantic properties of words are represented using their distributional patterns of co-occurrence in large volumes of text. Specifically, each word w is represented by a k-element vector reflecting the local distributional context of w relative to k context words $c_{1..k}$ and window size s. The value of each vector element is a function of the number of times each c occurs within s words of w in continuous text. Since k context words define a k-dimensional space, any two words w_1 and w_2 can be located by treating their vector representations as points in k-space. Distance measures in k space reflect the distributional similarity between w_1 and w_2.

[1] The normative association strength between two words w_1 and w_2 is measured by the percentage of subjects who produce w_2 as the first word that comes to mind when presented with w_1 in a free association task.

Distance in k-space can also be understood as a continuous measure of how easily two words can be substituted in the same context.

The assumption that the psychological concept of semantic relatedness corresponds to distributional similarity, and therefore to proximity in high-dimensional co-occurrence space, has proved useful to research in computational linguistics (e.g. Grefenstette, 1994; Schütze, 1993) and is becoming increasingly fruitful for psycholinguistic modelling (Landauer & Dumais, 1997; Levy, Bullinaria & Patel, 1997; Lund et al., 1995). However, if corpus-based models are to offer an adequate explanation of the representations underlying word recognition, they must be able to replicate the full range of priming effects found with human subjects. Simulations should reflect the range of lexical relations that have been shown to support priming, and also demonstrate the additive effect of normative association strength on lexical decision facilitation.

In this paper we attempt to replicate computationally each of the effects reported by Moss et al. in a single model using a corpus-based Contextual Similarity measure. In Experiment 1 (below) we present the model and examine the similarities and differences between human performance and our results using the Contextual Similarity measure. We then present a computational explanation for the additive effect of normative association, which is tested in Experiment 2.

Experiment 1

If corpus-derived similarity measures are to account for semantic priming effects (as shown by Lund et al., 1995 for category coordinates), a crucial test of the approach will be to see how well they account for priming between functionally related stimuli, as well as the 'associative boost' observed by Moss et al. These effects are the focus of Experiment 1, which attempts to replicate Moss et al.'s Experiment 2. This experiment was a speeded auditory lexical decision task with single word presentation of prime and target words.

Materials and Design

The design was identical to the original experiment, which varied three factors: Association, Semantic Type and Relatedness. Subtype was nested under Semantic Type: half of the Category Coordinates were natural objects (e.g. $\langle$dog, cat$\rangle$) and half were artifacts (e.g. $\langle$aeroplane, train$\rangle$). Correspondingly, the Functional semantic type was divided into words found in instrument relations (e.g. $\langle$knife, bread$\rangle$) and those found in scripts (e.g. $\langle$circus, lion$\rangle$).

Target words and their related primes were taken from Moss et al. (1995, Appendix 1).

Procedure

Co-occurrence vectors for each of the stimuli were extracted from approximately ten million words of the written portion of the British National Corpus (Burnage & Dunlop, 1992). The corpus was lemmatized by replacing inflectional variants with their canonical form from the CELEX database (Baayen, Piepenbrock & Van Rijn, 1993).

We chose a window size of ± 3 words since this corresponds roughly to the size of the phonological loop (Baddeley, 1990; Huckle, 1996, for further discussion).

Table 1: Difference in mean Contextual Similarity (Diff) versus amount of priming in milliseconds (Priming) for Associated prime-target pairs in each condition.

Condition	Related	Unrel	Diff	Priming
Cat Coord (all)	0.4036	0.1327	0.2709	94
Natural	0.4025	0.1342	0.2683	109
Artifact	0.4046	0.1311	0.2735	78
Functional (all)	0.2249	0.1147	0.1102	71
Script	0.2294	0.1182	0.1112	80
Instrument	0.2204	0.1113	0.1091	60

Table 2: Difference in mean Contextual Similarity (Diff) versus amount of priming in milliseconds (Priming) for Non-associated prime-target pairs in each condition.

Condition	Related	Unrel	Diff	Priming
Cat Coord (all)	0.3211	0.1393	0.1818	36
Natural	0.3261	0.1425	0.1836	57
Artifact	0.3160	0.1361	0.1799	16
Functional (all)	0.1630	0.1164	0.0466	41
Script	0.1636	0.0903	0.0733	31
Instrument	0.1623	0.1425	0.0198	50

Word vectors were created by advancing the window through the corpus. For each target word we recorded the number of times each of 446 context words[2] occurred within the window. Each element of the resulting vector was then replaced with its log-likelihood value, which can be considered as an estimate of how 'surprising' or distinctive a co-occurrence pair is (Dunning, 1993).

Several of the stimuli turned out to have an extremely low corpus frequency, which meant that vectors created from these words were likely to be unreliable. Consequently, we removed each prime-target pair that contained a word with an absolute lexeme frequency of less than 25, and balanced the other conditions by removing their lowest frequency pairs, leaving 12 items in each cell.[3]

We calculated the Contextual Similarity between each member of the Related prime-target pairs as the cosine of the angle between the vectors representing each word. Contextual Similarity for the Unrelated prime-target pairs was calculated as the mean of the Contextual Similarity of the target with each of the 11 other primes in the condition.

[2] The context words are those words estimated to be the most reliable in the corpus. For details of how reliability was estimated, see McDonald (1997). We excluded function words from consideration since we were interested in a model sensitive to semantic context, rather than to syntax, which is essentially captured by the set of function words (Finch & Chater, 1992). For a review of the linguistic and psychological evidence distinguishing function from content words, see Cann (1996).

[3] This procedure meant that the targets in the Associated and Non-associated conditions were no longer strictly matched for frequency. However, the Contextual Similarity measure we used is insensitive to the absolute values of the vector components.

Results

We carried out a three-way Analysis of Variance (Association $\times$ Semantic Type $\times$ Relatedness) on the Contextual Similarity measurements, with the Natural and Artifact and the Instrument and Script subtypes collapsed into the Category Coordinate and Functional semantic types, respectively. Contextual Similarity measurements and the corresponding reaction time data are summarised in Tables 1 and 2.

The simulation results proved to be very similar to those found in the original experiment: There was a main effect of Relatedness, $F(1, 92) = 73.61, p < 0.001$, indicating that collapsing over all conditions, semantically related prime-target pairs were more contextually similar than unrelated prime-target combinations (Contextual Similarity of 0.2781 and 0.1258, respectively).

We found a main effect of Semantic Type; Category Coordinates were significantly more contextually similar than materials in the Functional condition, $F(1, 92) = 23.25, p < 0.001$. There was also an interaction between Semantic Type and Relatedness, $F(1, 92) = 17.36, p < 0.001$. From Tables 1 and 2, it is clear that the Relatedness effect size is larger for Category Coordinates than for Functional items. These results differ from Moss et al., who found no reliable difference in the priming effect size between Category Coordinates and Functionally related pairs.

There was a significant interaction between Association and Relatedness, $F(1, 92) = 4.63, p < 0.05$. The Relatedness effect was larger for Associated than for Non-associated pairs. ANOVAs on the separated Associated and Non-associated conditions revealed significant Relatedness effects, $F(1, 46) = 40.22, p < 0.001$ and $F(1, 46) = 36.35, p < 0.001$, respectively, which correspond to the human results. This replicates the associative boost.

Consistent with the original experiment, the simulation failed to show an interaction between Association and Semantic Type, $F(1, 92) < 1$, and there was no three-way interaction between Association, Semantic Type and Relatedness, $F(1, 92) < 1$.

Since the three-way ANOVA revealed a significant difference between the Category Coordinate and Functional semantic types, we ran ANOVAs on each condition separately, in order to examine the relationship between Contextual Similarity and the type of relation more closely.

First, we carried out an ANOVA on the Functional materials. Contextual Similarity for Related word pairs was significantly larger than for Unrelated pairs, $F(1, 44) = 23.73, p < 0.001$. We found no main effects of Subtype $F(1, 44) < 1$ or Association, $F(1, 44) = 2.207, p = 0.14$. There were also no reliable interactions between the combinations of factors: Association x Subtype, $F(1, 44) < 1$; Subtype $\times$ Relatedness, $F(1, 44) < 1$; Association $\times$ Relatedness $\times$ Subtype, $F(1, 44) < 1$. The interaction between Association and Relatedness was marginally significant, however, $F(1, 44) = 3.91, p = 0.054$. Nearly identical results were obtained for the separated Category Coordinates.

Discussion

In summary, the pattern of results largely corresponded to those of Moss et al.'s Experiment 2. The significant difference in Contextual Similarity between semantically related and unrelated prime-target pairs replicates the overall priming effect found with human subjects. We also find an 'associative boost'; semantically related items that are also normatively associated are more contextually similar than non-associated materials.

Functional Relations vs. Category Coordinates

The main difference between the simulation and original results is the significant interaction between Semantic Type and Relatedness. Although separate ANOVAs for each condition verified a Relatedness effect for both Category Coordinates and Functional items in the simulation, it is clear that Contextual Similarity between Category Coordinate targets and their related primes is greater than for Functional prime-target pairs.

It is worth considering why this interaction should occur in a model constructed from co-occurrence statistics. We suggest that this is due to differences in the semantic roles typically filled by Category Coordinate and Functional items: there will be greater Contextual Similarity between 'bread' and its Category Coordinates such as 'fruit' and 'rolls' principally because Category Coordinates tend to fill the same position in predicate-argument structure *i.e.* the patient role associated with verbs such as 'serve' and 'eat'. Category Coordinates are therefore highly substitutable in context.

In contrast, items in Functional relationships tend to fill different semantic roles, such as instrument and patient, and will occupy different positions in predicate-argument structure. Moss et al.'s Instrument materials are a clear example of this — words were chosen that to fit the template:

you use a <prime> to do <target>.

In fact, functionally related prime-target pairs almost never occur in this way in real text. However, the vector representation of each word is a superposition of many separate occurrences, so the vector reflects the template structure better than any individual context. Since the Contextual Similarity measure does not distinguish left from right context, functionally-related prime and target words will share less of each other's context, and will be less contextually similar than Category Coordinates. This is sufficient to produce the Semantic Type $\times$ Relatedness interaction.

The Associative Boost

Although both Moss et al. (1995, Experiment 2) and our computational simulation demonstrated an Associated $\times$ Relatedness interaction (or associative boost), in a similar experiment, Shelton and Martin (1992) found that priming did not occur with items that were only semantically related, without also being normatively associated. Lund et al. (1995, Experiment 2) attempted to replicate this experiment using a corpus-based model similar to ours. But their simulation results were also incompatible with those of Shelton and Martin since they showed a simulated priming effect for the Semantic-only materials.

Lund et al. suggested that Shelton and Martin's Semantic materials were in fact less semantically related than their Associated stimuli, pointing out that 'semantic distances' derived from their model for word pairs in the Associated condition were smaller than for word pairs in the Semantic condition (although this difference only approached significance

$p = 0.061$). This reasoning is problematic, however, since distances in the model are assumed to reflect semantic relatedness in part *because* of their accord with semantic priming data.

Lund et al. (1995, Experiment 3) investigated this discrepancy by using stimuli from Chiarello, Burgess, Richards and Pollock (1990), which were more carefully controlled. In a simulation using these materials, they failed to find a interaction between Type of Relation (Associated, Semantic or Semantic+Associated) and Relatedness. A separate analysis of the Associated condition also failed to reveal a Relatedness effect. This result was in accord with data from human subjects (Lund et al., 1995, Experiment 4), confirming their hypothesis that a semantic relationship between prime and target words was necessary to induce a priming effect. However, results from another (human) lexical decision experiment (Lund et al., 1996, Experiment 1) using a new set of unrelated pairs, revealed a reliable effect for the Associated condition. There was no corresponding effect in their corpus-based simulation using this new stimuli set, however.

The lack of an interaction between Type of Relation and Relatedness in each of these priming simulations is inconsistent with the associative boost reported by Moss et al., which is also found in our replication. These differences warrant further discussion.

Moss et al. argue that the associative boost is due to priming at a different level of representation than the semantic level. They suggest that associative priming is a mechanism that is dependent on relationships between lexical forms. This is consistent with neural network models of priming (e.g. Moss, Hare, Day & Tyler, 1994; Plaut, 1995) that treat semantic and associative priming as due to fundamentally different types of information — associative priming effects result from contiguity between training items.

In contrast, we suggest that there is no need to treat semantic and associative relations differently, whether as separate levels of representation or distinct mechanisms. It is possible to account for semantic priming and the additive effect of normative association in the same corpus-based model. Specifically, the 'associative boost' falls out naturally from the way that co-occurrence statistics were compiled in the present experiment.

Explaining the Associative Boost

In order to address the difference between Lund et al.'s (1995; 1996) results and ours regarding the presence of an additive effect of association, it is necessary to examine the models' parameter settings in more detail.

One important difference is the window size within which co-occurrences are recorded. For example, if the Associated Coordinates 'cup' and 'saucer' nearly always appear within a ± 3 word window in the corpus, their Contextual Similarity will be high compared to prime-target pairs which merely tend to occur in similar (but not overlapping) contexts. If Associated pairs do frequently co-occur within the same window, their Contextual Similarity will be higher than matched Non-associated pairs, simply because of their shared context.

If Associated pairs tend to co-occur locally in a corpus of natural language, then the larger the window size, the more overlap of their immediate context, and hence the greater their Contextual Similarity. However, as the window size increases, the number of contextually irrelevant co-occurrences being recorded for each word grows, increasing the noisiness of the co-occurrence vectors, which reduces their Contextual Similarity. In order to verify this, we constructed versions of the model where the window size varied between ± 1 to ± 5 words. The best performance with respect to the human data was achieved with a window of ± 3 words.

We can now address Lund et al. (1995)'s claims about the Shelton and Martin materials. It may be true that several of the Associated pairs are more semantically related than items in the Semantic condition, though Shelton and Martin did attempt to control their materials for this variable through a relatedness rating questionnaire. But we suggest an alternative interpretation of Lund et al.'s observations: the Associated prime-target pairs are marginally closer together in vector space compared to the Semantic-only pairs because their co-occurrence vectors encode, in part, temporal co-occurrence as well as substitutability in context.

We suggest that Lund et al. did not find a reliable interaction between the Association and Relatedness factors in their Experiment 2 for the same reason that no associative boost was evident in the simulation results reported by Lund et al. (1995, Experiment 3; 1996, Experiment 1). The absence of this effect was due to the way the co-occurrence counts were collected. Although their window size was large (10 words), these simulations used the 200 most variant context words as vector components. This set will mostly consist of function words, which we suggest are simply not sufficiently specific indicators of semantic context.[4]

Our prediction about the origin of the associative boost in our semantic priming model can be easily tested: if the Contextual Similarity between two word vectors is affected by the temporal nature of language, we expect to find that the probability of lexical co-occurrence (in a ± 3 word window) to be greater for Associated Related pairs than for the Non-associated Related pairs taken from the materials for Experiment 1. This hypothesis is investigated in Experiment 2.

Experiment 2

Spence and Owens (1990) conducted a corpus-based investigation into the relationship between lexical co-occurrence and normative word association using the one million word Brown corpus. They found that associatively related word pairs tended to co-occur (within a window of 250 characters, or approximately 50 words) significantly more often in the corpus than pairs of words that were not normatively associated.

In Experiment 2, we test a similar hypothesis, that lexical co-occurrence in a window is more probable for the Associated materials in Experiment 1 than for Non-associated items. The null hypothesis is that the difference in co-occurrence probability between the Associated and Non-associated word pairs is not distinguishable from chance.

[4] We tried a similar approach with our corpus, using the 200 most variant lexemes occurring within a ± 3 word window as context words. In this case the Contextual Similarity differences between Associated and Non-associated pairs were negligible.

Procedure

Although the Associated and Non-associated targets were originally matched for median frequency, the same calculation over the lexeme frequencies in the ten million word corpus indicated that the median frequencies of the two conditions were not equal (57 per million vs. 38 per million). This discrepancy would bias simple lexical co-occurrence counts in favour of the Associated pairs, since if Associated target words occur more often in the corpus, they will have more chance of co-occurring with their corresponding prime words. Therefore, rather than comparing the raw co-occurrence counts of the prime-target pairs, we calculate conditional probabilities, which allows us to normalise for the target word frequency:

$$\mathsf{p}(\mathsf{prime} \mid \mathsf{target}) = \frac{\mathsf{f}(\mathsf{prime}\&\mathsf{target})}{\mathsf{f}(\mathsf{target})} \tag{1}$$

Results and Discussion

A Mann-Whitney U test revealed a highly significant difference between the lexical co-occurrence probabilities for Associated and Non-associated pairs ($U = 630, p < .00001$, one-tailed). The probability of an Associated target co-occurring with its corresponding prime was significantly higher than for the Non-associated pairs. We can therefore reject the null hypothesis, and conclude that differences in lexical co-occurrence probabilities for word pairs in these two conditions may be responsible for the difference in Contextual Similarity measurements because of the natural incorporation of local co-occurrence information into the context vectors.

The results of Experiment 2 are consistent with Spence and Owens' (1990) finding that lexical co-occurrence in a corpus and normative association strength are correlated: high associative strength predicts frequent lexical co-occurrence. McKoon and Ratcliff (1992) have additionally provided evidence that word pairs with a high probability of co-occurrence alone (*i.e.* not highly normatively associated) also show lexical decision facilitation. This result is also consistent with our model, although the amount of semantic relatedness present in their materials was not controlled.

Since the associative boost in semantic priming is modelled by the local co-occurrence probabilities in a large corpus, we hypothesise that the additive effect of normative association strength is better described as a variable subsumed by the more general phenomena of local co-occurrence. Our model would then predict an additive effect of high co-occurrence probability on priming between words that are semantically related.

Lund et al. (1996) have argued that normative association strength and lexical co-occurrence frequency are only correlated for the cases where a semantic relation is also present. The results of Experiment 2 are in agreement with this claim. Given that a semantic relationship holds between a prime-target pair, the prime is significantly more likely to occur within a small window of the target if the pair are also normatively associated.

General Discussion

In this paper we have demonstrated that corpus-based models of lexical semantics are capable of simulating the priming effects obtained with materials representing a wide range of lexical relations. In our computational simulation of a recent single word auditory lexical decision experiment (Moss et al., 1995, Experiment 2), we found that a corpus-derived measure of two words' distributional similarity corresponded well to the patterns of lexical decision facilitation observed in human subjects.

Experiment 1 showed that functional relations can be extracted from distributional statistics, but that the co-occurrence patterns of Category Coordinates and functionally related words are not equally informative. This was indicated by the interaction between Semantic Type and Relatedness in the simulation, which was not observed in the original experiment. We presented an explanation for this effect based on the idea that co-occurrence vectors encode coarse syntactic constraints as well as semantic regularities. Specifically, we suggested that Category Coordinates typically fill the same semantic roles at the level of predicate-argument structure, whereas Functional items tend to occupy different roles.

Experiment 1 also replicated the additive effect of normative association strength on the basic semantic priming effect observed by Moss et al. The interaction between Relatedness and Association, or 'associative boost', was modelled using the corpus-derived measure. The amount of lexical decision facilitation and Contextual Similarity between prime and target were greater for semantically related prime-target pairs that were also normatively associated, compared to pairs that were semantically related only.

We offered an explanation for why the associative boost occurs in the corpus-based model, and why previous research has not found this effect. We argued that the method used to collect co-occurrence statistics has a substantial influence on the 'associative' properties of the representations formed. In particular, associated word pairs are those that have a high probability of local co-occurrence. This hypothesis was confirmed in Experiment 2.

Conclusions

The adequacy of corpus-derived similarity measures for explaining the semantic priming effect was addressed by building on previous research (Lund et al., 1995) that investigated priming between category coordinates. We have demonstrated that the facilitation for category coordinates and functionally-related materials (prime-target pairs in instrument and script relations) found by Moss et al. (1995) can be modelled using the Contextual Similarity measure. Functional relations are often considered to be represented as extra-linguistic, schema-based or episodic knowledge. However, Moss et al. have shown that functional information can be accessed using the auditory lexical decision paradigm, even when stimuli are not normatively associated. The results of our computational replication in Experiment 1 have demonstrated that functional information is accessible directly from linguistic context, and that a significant part of the semantic priming effect can be attributed to the relationships between words, when defined in terms of their patterns of use.

The independent effects of association strength and semantic relatedness have been offered as evidence for distinct, qualitatively different priming mechanisms or representational levels. The results of the present experiments question the need for this distinction. A single level of representation can capture a wide range of lexical relations that support priming.

Acknowledgements

We would like to thank Richard Shillcock for commenting on earlier versions of this paper. Scott McDonald is supported by a Postgraduate Scholarship from the National Sciences and Engineering Research Council of Canada and the ORS Award Scheme. Will Lowe is supported by a Medical Research Council studentship.

References

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baddeley, A. D. (1990). *Human memory : Theory and practice*. Hove: Lawrence Erlbaum Associates.

Burnage, G. & Dunlop, D. (1992). Encoding the British National Corpus. In *Papers from the Thirteenth International Conference on English Language Research on Computerized Corpora*.

Cann, R. (1996). Categories, labels and types: Functional versus lexical. Edinburgh Occasional Papers in Linguistics EOPL-96-3, University of Edinburgh.

Chiarello, C., Burgess, C., Richards, L., & Pollock, A. (1990). Semantic and associative priming in the cerebral hemispheres: Some words do, some words don't . . . sometimes, some places. *Brain and Language*, *38*, 75–104.

Dunning, T. (1993). Accurate methods for the statistics for surprise and coincidence. *Computational Linguistics*, *19*, 61–74.

Finch, S. & Chater, N. (1992). Bootstrapping syntactic categories. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society*, (pp. 820–825). Bloomington, Indiana: Lawrence Erlbaum Associates.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory and Cognition*, *5*, 335–339.

Grefenstette, G. (1994). *Explorations in automatic thesaurus discovery*. Boston: Kluwer.

Huckle, C. C. (1996). *Unsupervised categorization of word meanings using statistical and neural network methods*. PhD thesis, Centre for Cognitive Science, University of Edinburgh.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of induction and representation of knowledge. *Psychological Review*, *104*, 211–240.

Levy, J., Bullinaria, J. A., & Patel, M. (1997). Using co-occurrence statistics for modelling language processes. Poster presented at the 3rd Conference on Architectures and Mechanisms for Language Processing (AMLAP), Edinburgh.

Lund, K., Burgess, C., & Atchley, R. A. (1995). Semantic and associative priming in high-dimensional semantic space. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, (pp. 660–665). Mahwah, NJ: Lawrence Erlbaum Associates.

Lund, K., Burgess, C., & Audet, C. (1996). Dissociating semantic and associative word relationships using high-dimensional semantic space. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, (pp. 603–608). Mahwah, NJ: Lawrence Erlbaum Associates.

McDonald, S. (1997). Exploring the validity of corpus-derived measures of semantic similarity. Paper presented at the 9th Annual CCS/HCRC Postgraduate Conference, University of Edinburgh.

McKoon, G. & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1155–1172.

McRae, K. & Boisvert, S. (1996). The importance of automatic semantic relatedness priming for distributed models of word meaning. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, (pp. 278–283). Mahwah, NJ: Lawrence Erlbaum Associates.

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science*, *6*, 413–427.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 863–883.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, (pp. 37–42). Mahwah, NJ: Lawrence Erlbaum Associates.

Schütze, H. (1993). Word space. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems*, volume 5. San Mateo, CA: Morgan Kaufmann.

Shelton, J. R. & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1191–1210.

Spence, D. P. & Owens, K. C. (1990). Lexical co-occurrence and association strength. *Journal of Psycholinguistic Research*, *19*, 317–330.