THE INFLUENCE OF SEMANTICS IN LEXICAL SELECTION IN SPEECH

R.A.I.DAVIES
Department of Experimental Psychology, School of Biological Sciences, Biology Building, Sussex University, Falmer, Brighton, BN1 9QG, U.K.
E-mail: radavies@biols.susx.ac.uk

The present chapter reports an experimental study of the semantic basis of lexical selection in speech. In a picture-word interference study, subjects named pictured objects while ignoring semantically related or unrelated distracters. It was found that the related distracters inhibited picture naming in comparison to the unrelated distracters. Indeed, the extent to which a related distracter elicited inhibition was predicted by the semantic similarity of distracter and target picture lexical concepts. Semantic inhibition is thought to show the effect of distracter presentation on the competition for lexical selection between candidate words in conceptually-driven speech. The results of the experiment suggest that such competition is more intense for more similar words. Measures of semantic similarity derived from lexical co-occurrence did not usefully predict observed semantic inhibition. The rated similarity of stimulus pairs was, however found to be useful in predicting inhibition. The results are discussed in relation to current models of lexical selection and of semantic space.

1 Introduction

When we speak we must quickly make an implicit choice between different words that variously match our intentions. The very large number of words available to a speaker and the very short time in which, often, he or she must choose the words make it likely that the choice is made in parallel. Such a claim is supported by observation of semantic substitution errors and blends in speech [33]. In a number of influential models of lexical retrieval in speech [10, 14 but see 29] it is assumed that information about the word needed to express an idea is retrieved in sequence, with first the word's semantic/syntactic specifications then its phonological form being accessed. It is assumed in these models that selection of the word takes place at the lexico-semantic or lemma level.

The present chapter reports a picture-word interference study investigating the nature of the semantic information that influences the choice we make every time we must express a concept in speech. In the picture-word interference task, the subject must name the picture while ignoring the interference stimulus, or distracter. It has been found that when the distracter is a word semantically related to the target name *e.g. horse-cat*, naming latencies are on average slower compared to when the distracter word is unrelated to the target name *e.g. horse-tree* [9, 12, 19, 27]. This delay, a semantic inhibition effect, has also been taken to reflect the parallel nature of lexical selection. It has been proposed that the distracter stimuli perturb a selection process that operates in the manner of a competition for selection between lexical candidates. Different candidates are activated in proportion to their match to the intentions of the speaker, the most activated candidate relative to all salient competitors is selected. In the case of semantically related competitors the level of competition is greater than it is for unrelated distracters because the related distracter is more similar to the target and so longer is required to distinguish the target word [27]. Opinion differs on whether the competition is implemented through lateral inhibition [30] or by means of a choice ratio [25] but for the purposes of the present study the essential point in our account is that semantic inhibition arises because candidates for selection in speech are compared simultaneously on the closeness of their fit to the intended meaning to be communicated.

Eliciting the semantic inhibition effect has appeared to depend upon the presentation of category co-ordinate distracters [27] but the effect, it has been claimed, is not influenced by the relative semantic similarity of target and distracter words[19]. Lupker [19] proposed that picture-word stimuli activate primarily category-level information, shared by target and distracter representations. If competition for selection depends only on category level information no effect of the relative similarity of target-distracter pairs will be detected. In one of the most influential current models of lexical selection [25], semantic interference arises as a result of automatic activation spreading between related nodes ('EXCLUDES' nodes between two mutually exclusive subordinates in a semantic category) in a network representation of semantic knowledge derived from Collins and Loftus' model [5]. The presentation of a related distracter primes activation of the target word but at the same time presentation of the picture primes the distracter. The greater activation of the distracter raises the level of competition, slowing the resolution of the selection process. Distracters do not facilitate picture naming because the priming of the distracter outweighs priming of the target. This is because, in the model, the links from target to distracter are shorter than those from distracter to target so that less activation is lost by the traversing of connections in the former case. In simulations reported by Roelofs [25], the weights on connections are not varied across items so that the model instantiates the implicit assumption that different degrees of relatedness do not influence the semantic inhibition observed. The related distracter slows naming more than the unrelated distracter because there are no links between the unrelated distracter and the target word so that the distracter does not benefit from target picture presentation. The model successfully simulated the semantic inhibition data recorded in a comprehensive study by Glaser and Dungelhoff [11].

It is the claim of the present study, however, that previous observations of the null effect of the relative similarity of target-distracter pairs have been flawed by invalid manipulations of similarity, warranting a reassessment of the influence of semantics on lexical selection. Lupker [19] varied the semantic similarity of target-distracter pairs in his study by varying their normative association or their typicality in addition to their co-ordinate status. He found that though distracters were both co-ordinate and highly associated with target words they did not inhibit naming any more than if they were simply co-ordinate. In an investigation reported by La Heij, Dirkx & Kramer [13], distracters that were both co-ordinates and strong associates of the target names facilitated picture naming at –400ms stimulus onset asynchrony (SOA) but inhibited naming at +75ms SOA[1]. La Heij et al. [13] interpret their results in terms of two different effects that have different time courses. The associated co-ordinate distracter facilitates naming at –400ms SOA (where the association effect predominates), has a null effect at 0ms (where co-ordinacy and association effects cancel) and inhibits at +75ms SOA (where co-ordinacy predominates). These findings undermine the view that variation in the association of picture-word pairs simply varies semantic similarity and therefore could (but do not) vary observed semantic inhibition. Association appears to be a different quantity to semantic similarity [18].

Lupker [13] also observed that typical co-ordinate distracters inhibited the naming of typical target names no more than did atypical co-ordinate distracters. Evidence reported by Cree, McRae & McNorgan [7] suggests that category typicality is not a useful measure of semantic similarity, certainly not to the extent that the latter influences lexical processing. Cree et al. [7] investigated the effect on lexical decision performance of prime words that were co-ordinates of the target word *e.g. squash* and were either more similar/less typical *e.g. pumpkin* or less similar/more typical *e.g. corn.* They found that mean decision latency was significantly facilitated for targets preceded by more similar/less typical primes. Less similar/more typical primes had no effect, however, in comparison to the baseline condition in which primes were unrelated. We can build expectations about speech performance on data from the primed lexical decision task if it is assumed that both picture-word semantic interference and lexical decision semantic priming stem from the same source, activation spreading across relatedness links at the semantic level of representation. Granted this assumption, it is clear that the manipulation of typicality in a semantic picture-word interference experiment is unlikely to influence the amount of semantic inhibition observed in such an experiment.

Further, evidence reported by McRae and colleagues [21,22] indicate that we can expect to find that semantic inhibition is observed in proportion to the semantic similarity of related target-distracter pairs. In a series of experiments, it was noted that when subjects were asked to categorise semantically printed target words categorisation performance was facilitated by semantically related prime words. Critically, the degree to which primes facilitated categorisation was predicted by the degree of semantic similarity obtaining between prime and target words. If again we are granted the assumption that we can apply findings from lexical decision or semantic categorisation tasks to form expectations about picture naming performance then we can hypothesise that the amount of semantic inhibition elicited by a related distracter will be predicted by its semantic similarity to the target picture name, beyond the mere co-ordinacy of the words.

The question now follows, how can one measure the semantic similarity of words? In the present study, target-distracter similarity was gauged by ratings of

---

[1] A negative asynchrony signifies that distracter stimulus onset precedes target onset. A positive asynchrony signifies that distracter onset follows target onset.

similarity and also measures of similarity obtained from Semantic Space Models (SSMs)[2] derived from the distributional analysis of lexical co-occurrence. McDonald ([20] p.13) comments that: "Word meaning varies along many dimensions; [Semantic Space Models] attempt to capture this variation in a coherent way, by locating words in a geometric space...words that are similar in meaning should be positioned closer together than words that are dissimilar..." The methods used to construct SSMs involve passing a 'window' over a speech or text corpus and counting how often a particular word occurs together with context words both before and after it within the window. The context words are a subset of the total number of word types in a corpus, usually limited to relatively frequent words. The obtained frequency or likelihood of co-occurrence between each word and every other context word are collated in a co-occurrence matrix after the analysis 'window' has been passed through the entire corpus. One can suppose that each context word constitutes a dimension on which every other word can be located in terms of their co-occurrence, thus defining a high-dimensional context space. For a particular word, the co-occurrence values in relation to other context words are the elements of a vector representing the word's co-ordinates in high-dimensional context space. We can evaluate the similarity of two words by calculating the distance between the points defined by their vectors.

In the present study, similarity data from two different models were used, the Lund and Burgess model (LBM)[3] [17] and McDonald's model (MM) [20]. These models share broadly the same methods of construction, as described above, but differ on several points. Most importantly, perhaps, LBM is built on an analysis of the co-occurrence of 70,000 words in an English language corpus of 320 million words of usenet text. Usenet text was chosen because it is claimed to provide a ready and plentiful supply of utterances. MM [20] is built on the analysis of the co-occurrence of 446 words in a corpus of 10.3 million spoken words, part of the British National Corpus [3]. In addition, LBM was constructed using a ten word window in which to count co-occurrences but MM [20] used a window of six words, three words before and three after the target word, and derived not co-occurrence frequencies but rather statistics of the log-likelihood of co-occurrence.

The applicability of such models to questions about the semantics of lexical selection in speech can be argued theoretically and on the basis of empirical evidence. Firstly, we can see that across an entire corpus each word may be expected to occur in contexts appropriate to its conventional meaning so that the co-occurrence statistics abstracted from a corpus can be taken as an index of usage which resembles or stands for an index of meaning. The similarity of the contexts in which two words tend to appear indicates how well one can be substituted for the

---

2 SSMs have been constructed from speaker-generated features [22] as well as from the analysis of lexical co-occurrence [17, 20] but the term Semantic Space Model will be used here to refer to the latter type of model only.

3 The acronym LBM is preferred here to the term HAL that Lund and Burgess themselves use because, as an anonymous reviewer points out, the term HAL has also been used in reference to an earlier computer system.

other in such contexts. If one word can easily be substituted for another it seems likely that both will compete for selection in speech when one is required to express a concept. Secondly, the SSM similarity of words has been found to significantly predict experimental effects or data. Lund, Burgess & Atchley [18] observed that the context similarity of related prime-target pairs was greater than that of unrelated pairs, when the prime words had been used successfully to elicit priming in human lexical decision performance. In addition, McDonald [20] compared rated similarity with context similarity of word pairs. He found that there was a highly significant correlation between the two measures of similarity. Despite these arguments it may still be the case that SSM similarity does not capture the kind of semantic similarity that influences lexical selection. It is a further hypothesis of the present study, therefore, that Semantic Space Models adequately capture the semantic similarity of target-distracter pairs eliciting semantic inhibition in a picture-word interference task.

1.1 Hypotheses

- Semantic inhibition observed in a picture-word interference task is related to the degree of semantic similarity between target and distracter.
- Semantic similarity is most adequately captured by the pair-wise similarities yielded by analysis of lexical co-occurrence.

2 Methods

2.1 Participants

Forty students of the University of Newcastle-upon-Tyne volunteered to participate in the experiment. All subjects had normal or corrected-to-normal vision. All were native speakers of English. All were female with an average age of 21.6 years (ranging from 18 to 49 years).

2.2 Materials

The target picture stimuli consisted of ninety-two line drawings of common objects. Eighty items were taken from the Snodgrass & Vanderwart [28] set of standardized pictures. The remaining items were taken from other sources ([8], Corel Draw 8 Clipart images, 1988-1997, Corel Corporation) but were drawn or were adapted so that their appearance was in the same style as the items from the Snodgrass and Vanderwart set.

Construction of the stimulus set focused on the target-distracter pairings. The manipulation of distracter type was within-items so that for the same target picture both a related and an unrelated distracter word was chosen. Related pairs were chosen first, as described below, then for each target the unrelated distracter was chosen so as to match the related distracter on syllable length, log lemma frequency and imageability, where data were available[4]. The unrelated distracter was chosen to be neither semantically nor phonologically related to the target picture name.

The principal aim in selecting the related target-distracter pairs was to choose a set consisting of category co-ordinates that spanned a wide range in the degree of semantic similarity per pair. An initial set of around two hundred candidate pairs was constructed. From this set, any pairs that were substantially associated were removed. In addition, pairs were not used if they were linked by a script relation e.g. restaurant-wine or by an instrument relation e.g. broom-floor [23]. Nor were they used if target and distracter were category co-ordinates at different levels of naming, say, superordinate (category name) and coordinate (basic level name) or coordinate-subordinate (see [26], [11], [32]).

For each remaining candidate related pair a set of values were obtained for the semantic similarity of target and distracter concepts. Measures of semantic similarity from two different semantic space models, LBM [17] and MM [20] were used. Pair-wise similarity calculated in terms of Euclidean and cityblock distance was derived from LBM. Similarity in terms of vector cosine values were derived from MM.

The cumulative distribution of the similarity of all candidates was calculated and pairs falling below the twenty-fifth percentile of the distribution were excluded from further use. Inspection of stimuli from a number of previous published studies of semantic inhibition had suggested an approximate cut-off point at that degree of similarity such that unrelated pairs tended to occur below it, related pairs above it.

The distribution of the pair-wise similarities of the final set of ninety-two related pairs was examined for each similarity measure. Both LBM distributions, Euclidean and cityblock distances, were reasonably normal. The cosine distribution was highly skewed toward low similarity values. The application of a number of transformations (such as taking the natural log or the log base ten of values) corrected this tendency but failed to render the cosine distribution normal by the light of statistical tests of skew [31]. In forthcoming correlation analyses, however, corrective transformations did not alter the overall pattern of correlations and therefore uncorrected cosine values continued to be used in all further analyses.

Ratings of pair-wise similarity were gathered following the final selection of target-distracter pairs. The pairs were presented to subjects in booklets in pseudo-random order such that no target or distracter was repeated in sequence. Subjects were asked to rate each pair on a seven point scale. Ten subjects were asked to rate the sensory similarity of target and distracter concepts. A different group of ten subjects were asked to rate their nonsensory or functional similarity; they were asked, "Do these objects do similar things? Are they found in similar situations?

4 CELEX [1]) frequency and syllable length data were available for all 92 related and unrelated distracters, but imageability data [6] were available for only 73 related distracters; unrelated distracters were matched to the average imageability of the related distracters where data were absent.

Or, are these objects used for similar things? Are they used in similar situations?" None of the subjects asked to provide ratings participated in the naming study.

Effort was made to ensure that the number of related pairs consisting of man-made or artefact objects approximately equalled the number of pairs consisting of living things.

2.3 *Apparatus*

The target pictures were scanned from print originals and digitized as bitmaps. Distracter stimuli were spoken by a native male speaker of English and recorded for auditory presentation as .wav files in a sound-attenuated studio. Stimuli were presented and responses were recorded by means of a Pentium III Windows 98 computer, using the DMASTR software developed at Monash University and at the University of Arizona by K.I. Forster and J.C.Forster (DMDX version 1.2.02). Subjects were seated in front of a computer screen at a distance of approximately 60cm so that on average a picture subtended 6.79 by 8.09 degrees of visual angle. Responses were recorded and stored as .wav files. Response latencies were extracted from these files 'by hand' using the Speech Station 2 speech analysis computer application.

2.4 *Design and Procedure*

The experimental design included stimulus onset asynchrony as a between-subjects and within-items factor with two levels (-100 and 0ms). It included distracter type as a within-subjects and within-items factor with three levels (silent no distracter, related distracter, unrelated distracter). Different groups of subjects performed the naming task at each SOA. All subjects were presented with all distracter types. All target pictures were used in all distracter conditions.

Each subject was individually tested in a dimly lit room. The experiment consisted of three parts. In the first, the familiarization phase, critical and practice pictures were presented to subjects together with their target names. In the second part of the experiment, the practice phase, subjects were given a block of eight trials in which to name pictures that were not used in the critical part of the test. The practice pictures matched the critical picture set on name length and frequency. In four practice trials, target pictures were accompanied by a distracter. In the other four trials, the pictures were presented alone. In the final part, subjects faced the critical picture naming trials. Subjects were instructed to name the pictures, ignoring the auditory distracter. There were 276 critical trials consisting of 92 pictures for naming each presented three times, each time with a different distracter type.

3 Results

Responses were classified as errors if: (1.) The name given did not match the target name. (2.) The response had been interrupted by a non-speech sound, or had been restarted or repaired. (3.) No response had been produced within the allotted response interval. (4.) A response had not been recorded due to mechanical failure. Under these error criteria, seven items were found to elicit incorrect responses at a rate greater than 10% of the responses made by all subjects to each item (coat, desk, nail, envelope, pigeon, scoop, spider); data pertaining to these items were excluded from all further analyses. This left a total of 10,200 data points, of which 218 or 2.18% were errors. No analysis of the errors is reported since the proportion of errors is so low.

3.1 Description of reaction time data

Table 1 shows the mean reaction times recorded in each condition. There is an effect of distracter type on naming performance. Response latencies were fastest under the 'silence' condition. The presentation of lexical distracters slowed latencies by approximately 100ms. Most critically, latencies in the related distracter condition were 14-20ms slower than latencies observed in the unrelated condition. This suggests that there was an effect, though it was relatively small, of the semantic relatedness of target-distracter pairings. It can be seen also that there was an effect of SOA. Reaction times were slightly faster when distracter stimulus onsets preceded picture onsets by 100ms. This SOA effect does not appear to influence the distracter effect, however.

Multiple regression analyses of the mean reaction times for each item are reported in the following section.

Table 1. Table showing the mean reaction times (ms) in each condition

	Stimulus onset asynchrony			
	0ms		-100ms	
Distracter condition	Reaction time (ms)	Standard deviation	Reaction time (ms)	Standard deviation
silence	646	59	630	83
unrelated	746	85	736	109
related	760	79	755	106

Note that figures indicate the mean in each condition of the average latency of each subject in each condition.

3.2 Multiple regression analyses

The dependent variable in the analyses was the mean latencies per item recorded in the related distracter condition. The independent variables were the mean unrelated latencies per item, plus the various measures of the semantic similarity of each related target-distracter pairs: LBM Euclidean or cityblock similarity; MM cosine similarity; rated sensory similarity; rated nonsensory similarity. Separate analyses were conducted on the data gathered in each SOA condition, allowing an indication of the validity of the regression solution across two different experimental settings. The regression analysis was conducted in an hierarchical fashion. Related latencies were regressed firstly on unrelated latencies, then on measures of semantic similarity. (In an analysis of the data gathered in a particular SOA condition, the related latencies in that condition was regressed on the corresponding unrelated latencies observed in that condition.) This method first accounts for all variation in the related latencies that can be ascribed to the effect of having an unrelated distracter. If there is a reliable effect of semantic relatedness it is represented together with error by the remaining variation left unaccounted. If that variation is adequately described by a measure of semantic similarity than that measure will be found to be a useful predictor of the variation in related latencies. All variables were converted to standard scores prior to analysis.

Related latencies correlated significantly with unrelated latencies (at –100ms SOA, $r = .53$, $p < .001$; at 0ms SOA, $r = .47$, $p < .001$). Related latencies correlated significantly, also, with rated nonsensory similarity (at –100ms SOA, $r = .23$, $p = .03$; at 0ms SOA, $r = .3$, $p = .005$) but with no other measure of pair-wise semantic similarity at the .05 significance level; the correlations between related latencies and rated sensory similarity merely approached significance (at –100ms SOA, $r = .2$, $p = .07$; at 0ms SOA, $r = .18$, $p = .1$). The LBM cityblock similarity of related pairs correlated significantly with the LBM Euclidean ($r = .83$, $p < .001$), MM cosine ($r = -.24$, $p = .04$) and rated sensory measures of similarity ($r = .22$, $p = .05$). LBM Euclidean similarity also correlated significantly with MM cosine ($r = -.35$, $p = .002$). The rated sensory and nonsensory similarity measures correlated significantly ($r = .63$, $p < .001$). The existence of substantial correlations between the different measures of semantic similarity warrants the adoption of the hierarchical rather than a simultaneous method of entering predictor variables in the regression analysis [2].

In a regression analysis of by-items data (data averaged across subjects for each item) the following variables were entered as predictors, in succession: first, unrelated latencies; second, LBM cityblock; third, LBM Euclidean; fourth, MM cosine; fifth, rated sensory similarity; sixth, rated nonsensory similarity. Table 3 presents statistics of R^2 change for the analysis.

Statistics of R^2 change allow us to see how useful a predictor variable is when it is used to predict the dependent variable alongside other predictors already included in the regression equation. It can be seen that the SSM measures of related pair-wise similarity add nothing significant to the capacity of the regression to predict the variation in related latencies. In contrast, R^2 is substantially increased by the entry to the model of rated sensory similarity. The high correlation between rated sensory and nonsensory similarity accounts for the fact that nonsensory similarity does not appear to make an independent contribution to the regression solution. (When nonsensory similarity is entered before sensory similarity it contributes significantly to the regression but sensory similarity does not.)

Principal Components Analysis (PCA) was used to extract the underlying factor that explains the correlation between sensory and nonsensory similarity. PCA indicated that 81.4% of the variance of each measure of rated similarity can be explained by a single common factor. In all further analyses the common factor is employed as the measure of rated similarity.

Table 3. Table showing the changing capacity of the regression to account for variation in the dependent variable with the addition of each predictor.

minus 100ms SOA

model	predictor variables included	R2	adjusted R2	R2 change	F change	df1	df2	F change significance
1	unrelated latencies at -100ms	0.20	0.19	0.20	18.98	1	74	< .001
2	unrelated RTs, cityblock	0.21	0.18	0.00	0.09	1	73	0.77
3	unrelated RTs, cityblock, Euclidean	0.21	0.17	0.00	0.01	1	72	0.94
4	unrelated RTs, cityblock, Euclidean, cosine	0.21	0.16	0.00	0.07	1	71	0.79
5	unrelated RTs, cityblock, Euclidean, cosine, rated sensory	0.30	0.25	0.10	9.46	1	70	0.003
6	unrelated RTs, cityblock, Euclidean, cosine, rated sensory, nonsensory	0.30	0.24	0.00	0.01	1	69	0.93

0ms SOA

model	predictor variables included	R2	adjusted R2	R2 change	F change	df1	df2	F change significance
1	unrelated latencies at 0ms	0.20	0.19	0.20	18.08	1	74	< .001
2	unrelated RTs, cityblock	0.20	0.17	0.00	0.01	1	73	0.93
3	unrelated RTs, cityblock, Euclidean	0.21	0.17	0.01	0.98	1	72	0.33
4	unrelated RTs, cityblock, Euclidean, cosine	0.21	0.16	0.00	0.16	1	71	0.69
5	unrelated RTs, cityblock, Euclidean, cosine, rated sensory	0.26	0.21	0.05	4.75	1	70	0.03
6	unrelated RTs, cityblock, Euclidean, cosine, rated sensory, nonsensory	0.28	0.21	0.02	1.66	1	69	0.20

In analyses of by-items data, related latencies (recorded in each SOA) were regressed as the dependent variable on the two predictors shown to be useful in the foregoing: unrelated latencies plus an extracted factor of rated similarity. It was found that regression solutions including both unrelated latencies and rated similarity significantly predicted the variation in related latencies at both SOAs (at –100ms SOA, adj. $R^2 = .331$, $F(2,82) = 21.8$, $P < .001$; at 0ms SOA, adj. $R^2 = .26$, $F(2,82) = 15.58$, $p < .001$). Both predictors making unique contributions. For unrelated latencies: at –100ms, $\beta = .54$, $p < .001$; at 0ms, $\beta = .45$, $p < .001$. For rated similarity: $\beta = .26$, $p = .005$; at 0ms, $\beta = .23$, $p = .02$).

It has been argued [16] that if one conducts a regression analysis on by-items data the outcome of such a test can be generalized across items but not across subjects. One can take into account variability in the value of regression coefficients across subjects by performing regression analyses of the data gathered for each subject alone, extracting for each a set of coefficient values. The reliability of the regression solution can then be tested in terms of subjects through testing the hypothesis that the mean regression coefficients (averaged across subjects' analyses) are significantly different to zero. It was found that a regression solution including as predictors unrelated latencies and rated similarity predicted variation in related latencies for most subjects. t-tests on coefficient values for the unrelated latencies predictor showed they were significantly different to zero (at –100ms, mean β =

.27, s.e. = .03, $t = 8.81$, 2-tailed $p < .001$; at 0ms, mean $\beta = .26$, s.e. = .03, $t = 8.98$, 2-tailed $p < .001$). Likewise, t-tests on values for the rated similarity predictor were significant (at –100ms, mean $\beta = .15$, s.e. = .02, $t = 6.56$, 2-tailed $p < .001$; at 0ms, mean $\beta = .13$, s.e. = .02, $t = 6.49$, 2-tailed $p < .001$). It is plain that the regression solutions yielded in the by-items analyses are consistent in predicting the related naming latencies for all subjects participating in the experiment.

4 Discussion

The present study was designed to test two hypotheses. It was proposed that the semantic inhibition observed in a picture-word interference task is related to the degree of semantic similarity between target and distracter. In addition, it was proposed that that semantic similarity would be adequately described by measures yielded by two Semantic Space Models. The evidence gathered in the experiments support the first but not the second hypothesis. In regression analyses, it was found that naming latencies recorded in the related distracter condition were significantly predicted by the semantic similarity of pairs, when variation due to the presentation of unrelated distracters had also been accounted. The measure of similarity useful in predicting related latencies was the rated similarity of stimulus pairs. There was no evidence that SSM measures of similarity had any significant role.

The observation of a significant relationship between semantic inhibition and the semantic similarity of target and distracter words is contrary to the claim made by Lupker [19] that competition for selection is based only on shared category-level information. The results of the present study make it plain that, to the extent semantic inhibition reflects competition for lexical selection, words competing to be used in speech compete more when they are more similar. The data do not allow us to evaluate whether this means that competition for selection is between similar words or between words that are category co-ordinates, such that the level of competition between co-ordinates is greater if the words are more similar. At the least, however, it now appears that present models of lexical selection, such as Roelofs' [25] model, must be adapted to account for the semantic similarity effect.

The finding that rated similarity but not SSM similarity helps to predict semantic inhibition suggests that the SSMs do not adequately describe the semantic similarity involved in lexical selection in speech. What does this say about the adequacy of SSMs? One of the most impressive findings of recent years have been the demonstrations that significant semantic information can be abstracted from lexical co-occurrences. We are still however working out the applicability of this approach to the explanation of human behaviour.

LBM and MM cannot be claimed to represent the class of SSMs derived from lexical co-occurrence analyses, though LBM is amongst the most well known. As has been described there are a variety of parameters which specify how such models are constructed. It has been shown that differences in the size of the corpus used, in window size, window type, the number and type of counted context words substantially influence the performance of SSMs [24, 15]. Further, Patel and colleagues [24] observed that optimal values for parameters varied in relation to which performance criteria were adopted. It may be the case therefore that lexical co-occurrence derived SSMs based on different sets of parameter values could be shown to yield similarity measures that predict semantic inhibition.

Knowing that the performance of SSMs vary in relation to the variation of parameter values how can we find out whether SSMs can be used to describe the semantic information used in lexical selection in speech? A study that could evaluate the applicability of SSMs in general to the particular problem of lexical selection would have to ground null results, if any such were to be observed, on an adequate sampling of the space of possible SSMs. The study might be manageable if this space was limited in a principled fashion. One could do so by limiting the space of possible SSMs in terms of human empirical data, using window sizes that resemble the limits on working memory capacity [2], and using corpora similar in size and kind to recorded or estimated human lexical experience (as Patel and colleagues [24] suggest).

In conclusion, the present study has demonstrated that lexical competition for selection is likely to be based on more than just the information shared by category co-ordinates. The relationship between semantic similarity and semantic inhibition shows that lexical competition depends upon semantic content detailed to the level of individual concept similarities. This relationship was predicted by ratings of pair-wise similarity but not by measures of similarity yielded by lexical co-occurrence based SSMs. Further empirical work will clarify the adequacy of lexical context derived models of semantic space.

5 Acknowledgements

The work reported in this chapter was conducted during the author's PhD at the University of Newcastle upon Tyne, funded in part by the UK ESRC (Grant no. R00429724430). I would like to thank David Howard for comments and advice that have helped me in writing the chapter. I thank Curt Burgess and Scott McDonald for providing semantic similarity data from their models.

References

1. Baayen, H., Piepenbrock, R. and van Rijn, H., The CELEX lexical database, CD-ROM (Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA, 1993).
2. Burgess, C. and Lund, K., The dynamics of meaning in memory. In E. DIETRICH and A.B. MARKMAN (Eds.) *Cognitive dynamics: Conceptual and representational change in humans and machines.* (Lawrence Erlbaum Associates, Mahwah, NJ, 2000).

3. Burnage, G. and Dunlop, D., Encoding the British National Corpus. In *Papers from the 13th international conference on English language research on computerised corpora.*
4. Cohen, J. and Cohen, P., Applied multiple regression/correlation analysis for the behavioural sciences (2nd edition) (Lawrence Erlbaum Associates, Hillsdale NJ, 1983).
5. Collins, A.M. and Loftus, E., A spreading-activation theory of semantic processing, *Psychological Review* **82** (1975) pp. 407-428.
6. Coltheart, M., The MRC psycholinguistic database, *The Quarterly Journal of Experimental Psychology* **33A** (1981) pp. 497-505.
7. Cree, G.S., McRae, K. and McNorgan, C., An attractor model of lexical conceptual processing: Simulated semantic priming, *Cognitive Science* **23** (1999) pp. 371-414.
8. Cycowicz, Y.M., Friedman, D. and Rothstein, M., Picture naming by young children: norms for name agreement, familiarity, and visual complexity, *Journal of Experimental Child Psychology* **65** (1997) pp. 171-237.
9. Damian, M. and Martin, R.C., Semantic and phonological codes interact in single word production, *Journal of Experimental Psychology: Learning, Memory and Cognition* **25** (1999) pp. 345-361.
10. Dell, G.S., A spreading-activation theory of retrieval in speech production, *Psychological Review,* **93** (1986) pp. 283-321.
11. Glaser, W. and Dungelhoff, F.-J., The time course of picture-word interference, *Journal of experimental Psychology: Human Perception and Performance* **10** (1984) pp. 640-654.
12. La Heij, W., Components of Stroop-like interference in picture naming, *Memory and Cognition* **16** (1988) pp. 400-410.
13. La Heij, W., Dirkx, J. and Kramer, P., Categorical interference and associative priming in picture naming, *British Journal of Psychology* **81** (1990) pp. 511-525.
14. Levelt, W.J.M., Roelofs, A. and Meyer, A.S., A theory of lexical access in speech production, *Behavioural Brain Sciences* **22** pp.1-75.
15. Levy, J.P. and Bullinaria, J.A., Learning lexical properties from word usage patterns: Which context words should be used? In R.F.FRENCH and J.P.SOUGNE (Eds.) *Connectionist models of learning, development, and evolution: Proceedings of the Sixth Neural Computation and Psychology Workshop* (Springer, London, 2001) pp. 273-282.
16. Lorch, R.F. and Myers, J.L., Regression analyses of repeated measures data in cognitive research, *Journal of Experimental Psychology: Learning, Memory and Cognition* **16** (1990) pp. 149-157.
17. Lund, K. and Burgess, C., Producing high-dimensional semantic spaces from lexical co-occurrences, *Behaviour Research Methods, Instruments and Computers* **28** (1996) pp. 203-208.
18. Lund, K., Burgess, C. and Atchley, R.A., Semantic and associative priming in high-dimensional semantic space, *Proceedings of the Cognitive Science Society* (Lawrence Erlbaum Associates Inc., Hillsdale NJ, 1995) pp.660-665.

19. Lupker, S.J., The semantic nature of response competition in the picture-word interference task, *Memory and Cognition* 7 (1979) pp. 485-495.
20. McDonald, S., Environmental determinants of lexical processing effort, (PhD thesis, University of Edinburgh, 2000).
21. McRae, K. and Bosivert, S., Automatic semantic similarity priming, *Journal of Experimental Psychology: Learning, Memory and Cognition* **24** (1998) pp. 558-572.
22. McRae, K., Seidenberg, M. and de Sa, V.R. On the nature and scope of featural representations of word meaning, *Journal of Experimental Psychology: General* **126** (1997) pp. 99-130.
23. Moss, H., Ostrin, R.K., Tyler, L.K. and Marslen-Wilson, W.D. Accessing different types of lexical semantic information: Evidence from priming, *Journal of Experimental Psychology: Learning, Memory and Cognition* **21** (1995) pp. 863-883.
24. Patel, M., Bullinaria, J.A. and Levy, J.P., Extracting semantic representations from large text corpora, In J.A. BULLINARIA, D.W. GLASSPOOL and G. HOUGHTON (Eds.) *4th. Neural Computation and Psychology Workshop, London 9-11 April 1997: Connectionist Representations* (Springer, London, 1997) pp. 199-212.
25. Roelofs, A., A spreading-activation theory of lemma retrieval in speaking, *Cognition* **42** (1992) pp. 107-142.
26. Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M. and Boyes-Braem, P., Basic objects in natural categories, *Cognitive Psychology* **8** (1976) pp. 382-439.
27. Schriefers, H., Meyer, A.S. and Levelt, W.J.M., Exploring the time-course of lexical access in language production: picture-word interference studies, *Journal of Memory and Language* **29** (1990) pp. 86-102.
28. Snodgrass, J.G. and Vanderwart, M., A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity, *Journal of Experimental Psychology: Human Learning and Memory* **6** (1980) pp. 174-215.
29. Starreveld, P.A. and La Heij, W., Time-course analysis of semantic and orthographic context effects in picture naming, *Journal of Experimental Psychology: Learning, Memory and Cognition* **22** (1996) pp. 869-918.
30. Stemberger, J.P,. An interactive activation model of language production. In A.W. ELLIS (Ed.) *Progress in the psychology of language* (Erlbaum, London, 1985) pp. 143-186.
31. Tabachnick, B.G. and Fidell, L.S., *Using multivariate statistics* (4th. edition) (Allyn and Bacon, Needham Heights MA, 2001).
32. Vitkovitch, M. and Tyrell, L., The effects of distracter words on naming pictures at the subordinate level, *The Quarterly Journal of Experimental Psychology* **52A** (1999) pp. 905-926.
33. Wheeldon, L.R. and Monsell, S., Inhibition of spoken word production by priming a semantic competitor, *Journal of Memory and Language* **33** (1994) pp. 332-356.