Title
Dissociating Semantic and Associative Word Relationships Using High-Dimensional Semantic Space

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Authors
Lund, Kevin
Burgess, Curt
Audet, Chad

Publication Date
1996

Peer reviewed

Dissociating Semantic and Associative Word Relationships Using High-Dimensional Semantic Space

Kevin Lund
Department of Psychology
University of California, Riverside
Riverside, CA 92521
kevin@locutus.ucr.edu

Curt Burgess
Department of Psychology
University of California, Riverside
Riverside, CA 92521
curt@cassandra.ucr.edu

Chad Audet
Department of Psychology
University of California, Riverside
Riverside, CA 92521
chad@buffy.ucr.edu

Abstract

The Hyperspace Analogue to Language (HAL) model is a methodology for capturing semantics from a corpus by analysis of global co-occurrence. A priming experiment from Lund et al. (1995) which did not produce associative priming with humans or in the HAL simulation is repeated with rearranged control trials. Our experiment now finds associative priming with human subjects, while the HAL simulation again does not produce associative priming. Associative word norms are examined in relation to HAL's semantics in an attempt to illuminate the semantic bias of the model. Correlations with association norms are found in the temporal sequence of words within the corpus. When the associative norm data are split according to simulation semantic distances, a minority of the associative pairsthat are close semantic neighbors are found to be responsible for this correlation. This result suggests that most associative information is not carried by temporal word sequence in language. This methodology is found to be useful in separating typical "associative" stimuli into pure-associative and semantic-associative subsets. The notion that associativity can be characterized by temporal association in language receives little or no support from our corpus analysis and priming experiments. The extent that "word associations" can be characterized by temporal association seems to be more a function of semantic neighborhood which is a reflection of semantic similarity in HAL's vector representations.

One of the most robust cognitive phenomena is the lexical/semantic priming effect. The word-recognition literature is replete with extensions of Meyer and Schvaneveldt's (1971) initial finding that a word prime that is semantically related to a word target (CAT-DOG) will lower the recognition threshold for the target as compared to a semantically unrelated word pairing (FLOWER-DOG). In most lexical/semantic theories, the role of association is thought to be an important component to semantic structure (Collins & Quillian, 1969; McClelland & Kawamoto, 1986; Moss et al., 1995; Plaut, 1995). More recently, a number of investigators have presented evidence that the phenomenon known as the semantic priming effect is essentially associative in nature. Lupker (1984) showed that words that were associated, but not categorically related, produced a reliable priming effect using the naming task. However, words that were only categorically related did not show a priming effect. One conclusion that can be drawn from this research was that semantic relations are accessed later than the more basic associative relations. More recently, Shelton and Martin (1992) proposed that automatic retrieval of word information in priming is associative. They used a lexical-decision methodology, but tried to insure automatic processing by having subjects make a response to every item (rather than having obvious prime-target pairs) and by having the related trials form only a small proportion of the total trials. Their conclusion, similar to that of Lupker, was that only associatively related items resulted in automatic priming and that semantic (categorical) effects required some type of controlled process, perhaps subject expectancy. From a computational viewpoint, Plaut (1995) presented a model of priming showing that priming effects could be easily be demonstrated by temporal association and that semantic characteristics did not seem to contribute to this effect.

We think that it may be premature to conclude, however, that the mechanism underlying the priming effect is associative rather than semantic. There are two issues that have to be tackled prior to making any definitive conclusion. First, in word relationships, aspects of semantic and associative meaning are typically correlated. This confounding of associative and semantic relationships leads to confusion about how this information might be represented in memory (and in memory models) as well as about how stimuli should be selected for experiments. Thus, it is crucial in experiments and simulations dealing with this issue that there be a straightforward way to operationally define semantics and association, and, moreover, that this operational definition likewise have a straightforward correspondence to a range of empirical findings that already exist. The second issue that has to be adequately addressed is how can semantic representations be modeled in such a way that allows for experimentation across the broad range of semantic phenomena beyond just priming effects? Such phenomena could include basic categorization, the prediction of human semantic judgments (Lund & Burgess, in press), semantic effects in sentence comprehension (Burgess & Lund, 1996), semantic neighborhood effects, as well as semantic impairments that are associated with deep dyslexia and other types of brain damage (Buchanan, Burgess, & Lund, 1996; Burgess & Lund, in press).

The Hyperspace Analogue to Language (HAL) model of semantic memory offers such an approach. In previous papers, we have proposed that the representations acquired by HAL are semantic in nature and that these semantic representations are what underlie the basic priming effect (Lund et al., 1995; Lund & Burgess, in press). For example, experiment 2 of Lund et al. suggests that the priming effect obtained by Shelton and Martin (1992) for the associated word pairs was, at least in part, carried by the semantic similarity of the items. In addition, the absence of a priming effect for the items related only by semantic relatedness may have been due to insufficient semantic similarity.

In our first experiment, we present an extension of our earlier semantic and associative priming experiments that further supports our claim that semantic priming occurs as a function of semantic similarity rather than associative strength of prime-target pairs. In addition, these results suggest that the representations generated by HAL are semantic rather than associative in nature. Experiments 2 and 3 utilize word pairs from word association norms and two operational definitions of associativity. We conclude that semantic neighborhoods can be used to dissociate the semantic and associative components in word meaning.

Simulation Methods

Matrix Construction

The basic methodology of the simulation is to develop a matrix of word co-occurrence values for a given vocabulary. This matrix is then divided into co-occurrence vectors for each word, which can be analyzed for semantic content (see Lund and Burgess, in press, for a more complete account of the matrix-construction methodology and validating studies). An analysis of co-occurrence must define a window size, that is, the largest number of words occurring between a pair of words such that the pair may be considered to co-occur. The limiting case of a small (useful) window is a width of one, which would correspond to counting only immediately adjacent words as co-occurrants. At the other end of the spectrum, one may count all words within a logical division of the input text as co-occurring equally (see Landauer & Dumais, 1994; Schvaneveldt, 1990). A very small window may miss constructs spanning several words (such as lengthy noun phrases), while large windows risk introducing large numbers of extraneous co-occurrences. Therefore, we chose a window width of ten words.

Within this ten-word window, co-occurrence values are inversely proportional to the number of words separating a specific pair. A word pair separated by a nine-word gap, for instance, would gain a co-occurrence strength of one, while the same pair appearing adjacent to one another would receive an increment of ten. Cognitive plausibility was a constraint, and a ten-word window with decreasing co-occurrence strength seemed within these bounds (Gernsbacher, 1990). The product of this procedure is an N-by-N matrix, where N is the number of words in the vocabulary being considered.

Text Source. The corpus that was analyzed was approximately 160,000,000 words of English text gathered from Usenet. All newsgroups containing English dialog were included. This source was chosen both for volume and content. Roughly ten million words per day are available, with conversational content spanning a wide array of topics.

Vocabulary. The vocabulary used for the analysis consisted of the 70,000 most frequently occurring symbols within the corpus. A check against the standard Unix dictionary showed that about one half of these were valid English words. The remaining symbols were common misspellings, slang, proper names, and sequences of punctuation or numbers.

Similarity Measurement. Once the matrix was constructed, word vectors were extracted. Each word in the vocabulary has one row and one column to represent it in the matrix. By combining the data from a word's row and column, a vector of 140,000 elements is formed for that word.

These vectors may then be compared by any appropriate distance metric. The distance metric used in the following experiments is the Euclidean distance measure. As the Euclidean metric is sensitive to vector magnitude, the vectors are normalized to a constant length before being compared.

Conceptually, these vectors represent points in a high-dimensional space. In this case, 140,000 dimensions, although only a small subset of the vector elements are necessary to produce semantic effects. For example, 200 elements (i.e., a 200 dimensional space) were used in Lund and Burgess (in press); the dimensionality was reduced by retaining only the most variant portions of the co-occurrence matrix. Other methods of dimensionality reduction have also been used with success; Landauer and Dumais (1994), for instance, used singular value decomposition. With high-dimensional semantic space models, similarity is conceptualized as corresponding inversely to inter-point distances; i.e., presumably the more similar two words are, the closer their points.

Experiment 1

Lund, Burgess, and Atchley (1995) used stimuli representing three types of word relationships in a series of experiments in an attempt to dissociate semantic from associative sources of priming. Semantically related words (TABLE - BED) are instances of the same category and share a number of features. Associated words (MOLD - BREAD) are those which are associated as determined by human word association norms. The items selected, however, were not instances of the same category and, therefore, share few semantic features. The third type of word relation are pairs that are both semantically and associatively related (UNCLE - AUNT). These different word relations should allow one to distinguish between the associative and semantic components of priming. Lund et al. (1995) argued that the vector representations generated by HAL are semantic in nature and not associative. Their evidence for this was that by using semantic distances, priming (unrelated - related) was obtained for the semantic as well as for the semantic + associated pairs, but not for the pairs that were related only through their association (Exp. 3). Using the same stimuli pairs in a lexical decision experiment with human subjects, the identical pattern of results was found.

While the parallels between the human and simulation results were striking in the Lund et al. (1995) paper, the parallels between the human results and earlier research were not. Specifically, neither the simulation nor the human subjects showed reliable priming on stimuli which were related only by word associativity. This was somewhat troubling in that associative priming has been demonstrated in earlier research (Chiarello, Burgess, Richards, & Pollock, 1990; Fischler, 1977). Closer inspection of the reaction times in Lund et al. (1995, Exp. 4, see Table 1) shows that the reaction times for the related word pairs in the associated-only condition were actually faster than in the semantic-only case. The lack of a priming effect for the associated-only condition may be due to an unreliable, unrelated baseline condition. A similar pattern is seen in the simulation results. In order to clarify these results, a new set of unrelated word pairs was formed by rearranging the related word pairs; if the prior results occurred due to a problematic set of unrelated word pairs, associated priming should be obtained with human subjects. If, as Lund et al. (1995) claim, the representations generated by HAL are not associative, but semantic, the simulation results should still reflect a lack of an associative priming effect.

	HAL simulation			Human subjects		
	Sem.	Assoc.	Both	Sem.	Assoc.	Both
R	347	322	331	643	623	603
U	413	339	391	673	634	631
U-R	66	17	60	30	11	28
Std.	1.00	0.26	0.91	1.00	0.36	0.93

Table 1. Simulation and human results (from Lund, Burgess, & Atchley, 1995)

Methods

Twenty-two undergraduate students participated in order to earn course credit. The stimuli used in this lexical-decision priming experiment were taken from Chiarello et al. (1990) and consisted of word pairs of three relational types (associative, semantic, and semantic + associative). These related word pairs were rearranged to form unrelated word pairs; additionally, a number of word-nonword trials equal to the combined number of related and unrelated word-word trials were included. Target words were balanced for both word length and printed frequency, yielding a total of 288 word pairs.

An experimental list included all 288 trials and was preceded by four "warm up" trials. Word primes were counterbalanced so that a target would be preceded by a related word in one list and an unrelated word in a second list. This allowed the targets to act as their own controls. Of the related word-word pairs, one third were word pairs that were only semantically related, one third were only associatively related, and the remaining pairs were both semantically and associatively related.

Stimulus presentation and timing was conducted on PC clones. Each trial began with a 500 ms fixation cross, followed by a prime at this location for 300 ms immediately followed by the target which remained until either the subject made a lexical decision or 2500 ms elapsed. Accuracy feedback was provided, as well as a time-out signal for lack of response within 2500 ms. A set of fifteen practice trials was presented first.

Results and Discussion

Results are shown in Table 2. As desired, the variance of the unrelated condition for human subjects was reduced (from 11 to 6.2). Reaction times were faster for related trials than for unrelated, $F(1, 285) = 17.91$, $p < 0.0001$. There was an effect of associative type ($F(5, 281) = 2.6$, $p = 0.02$) but no type by relatedness interaction ($F < 1$). Planned comparisons were made at each level of word relation in order to determine priming by stimulus type. Priming was found for all three conditions: semantic, $F(1, 93) = 4.19$, $p = 0.04$; associated, $F(1, 94) = 8.11$, $p = 0.005$; and both semantic and associated, $F(1, 94) = 6.64$, $p = 0.011$. These results are consistent with our hypothesis that the earlier lack of associative priming with human subjects was due to a poor selection of unrelated word pairs. However, in both experiments, the standard procedure was followed to obtain unrelated word pairs (manual rearrangement, being careful to include no clearly related word pairs). The resulting sporadic associative priming in these experiments, contrasted with the consistency of semantic priming, leads us to conclude that the role of associative priming in a lexical decision task is not as straightforward as that of semantic similarity.

The HAL simulation results for the newly rearranged word pairs are similar to those obtained previously. There is an overall priming effect, $F(1, 285) = 21.02$, $p < 0.0001$. There was no reliable effect of associative type ($F < 1$), nor a reliable interaction ($F(2, 281) = 1.26$, $p = 0.28$). Priming is present for the semantic condition, $F(1, 93) = 5.92$, $p = 0.017$, and for the semantic plus associated condition, $F(1, 94) = 10.95$, $p < 0.0013$, but once again no priming is found for the associated-only condition, $F < 1$. Standardized scores are computed by dividing all priming amounts by the semantic priming amount, in order to ease comparisons of priming magnitudes between the simulation and the human results (see Table 2).

The simulation was not sensitive to the changes in stimuli which brought about human associative priming. This result provides further that HAL is more sensitive to semantics than to associativity.

	HAL simulation			Human subjects		
	Sem.	Assoc.	Both	Sem.	Assoc.	Both
R	540	469	416	615	583	591
U	513	487	494	652	626	629
U-R	37	18	78	37	43	38
Std.	1.00	0.48	2.11	1.00	1.16	1.02

Table 2. Simulation and human results

Experiment 2

Experiment 1 showed that the HAL simulation does not produce statistically reliable priming with associated-only word pairs. However, there is a consistent, although not reliable, relatedness advantage found for associated-only stimuli, both in the experiment presented here and in other research. Furthermore, the semantic + associated condition in Experiment 1 produced greater priming than did the semantic-only condition.

For these reasons, a blanket declaration of lack of associativity in HAL's representations seems premature. However, it does seems clear that whatever associative information is available in HAL's vectors will not be found though priming simulations. This is important to clarify since we want to claim that semantic similarity makes an important contribution in priming with humans.

Stimuli construction in word priming and other psycholinguistic experiments often employ word association norms. These norms provide a compilation of a set of people's frequency ordered associations to many words. Word relationships that can be found in word association norms can take a variety of forms. It is not unusual to find many types of associative and semantic relationships in these collections of items. These norms, then, provide an important test of the representations that can be generated using the HAL model. In this experiment, word pairs from the Palermo and Jenkins (1964) norms will be compared with two potential associativity indices available from the co-occurrence matrix. Additionally, the Palermo and Jenkins' data will be compared to semantic distances from the HAL simulation.

Methods

The two HAL-derived candidate associativity indices are co-location frequency and co-location separation. Co-location frequency, computable for any two words, is simply a raw count of the number of times that those two words occurred within the ten-word co-occurrence window during construction of the co-occurrence matrix. This corresponds to one intuitive definition of associativity, namely that words are associated to the degree that they tend to occur together in language (Miller, 1969; Spence & Owens, 1990). This measure is not scaled or normalized for word frequency.

The second associativity measure, co-location separation, is a measure of how many words, on average, separated a certain word pair as they occurred within the co-occurrence window. For instance, if the phrase "ladies and gentlemen" was the only context in which the words "ladies" and "gentlemen" ever occurred within the corpus, those two words would have a co-location separation measure of 1.0, as they are always separated by one word, "and." As the co-occurrence window used in these experiments was ten words wide, this measure can range from zero to nine. Co-location frequency and separation values for some example word pairs are give in Table 3.

These potential associativity metrics were compared to Palermo and Jenkins' (1964) word association norms. In their norms, target words were available along with associates produced by human subjects. Each target-associate pair was ranked by how many subjects produced that particular response. For each target word in the human associativity data, the top five associates were used to form five word pairs (e.g., man-boy). Co-location frequency and separation values for these word pairs were then computed, as well as semantic distances, and these values were examined for correlations.

Word Pair		Col. Sep.	Col. Freq.
black	white	2.9	3846
man	woman	3.4	1497
led	zeppelin	0.4	290
holy	cow	0.0	64
man	tall	5.2	49

Table 3. Example word pairs.

Results and Discussion

Correlations between human associativity ratings and both co-location separation and frequency are shown in Table 4. Consistent with our earlier claim, there was no reliable correlation between semantic distances and associativity ratings. There was a small correlation between co-location separation and associativity ($r = -0.1$, $p = 0.05$), indicating that, to a minor degree, more highly associated word pairs occur closer to each other than do less associated pairs.

By far the largest correlation found was that between co-location frequency (raw count of co-occurrence within a ten-word window) and word pair associativity ($r = 0.25$, $p < 0.0001$). This supports the intuitive and common operational definition of associativity that associated words tend to appear close to each other in language. It is notable, though, that the larger of the significant correlations was that between associativity and co-location frequency rather than between associativity and co-location distance. In other words, associativity for a word pair was not well predicted by how close, on average, the two words were to each other when they were within the ten-word window, but was predicted rather well simply by the number of times which they co-occurred within the window at any distance.

These results can illuminate much about HAL, semantics, and distributed semantic representations. An examination of the word pairs used as associative pairs reveals substantial semantic overlap. For instance, the top two associates of *man* are listed as *woman* and *boy*.

The associativity of the word pairs used in this experiment is not in doubt; clearly, *man* and *woman* are highly associated (all these stimuli were the strongest five associates to each target). However, a great many of the word pairs are both associated and semantically similar. Experiment 3 will examine this phenomenon in greater detail.

	Distance	Col. Sep.	Col. Freq.
Human Associativity	-0.06 ($p = 0.237$)	-0.1 ($p = 0.05$)	0.25 ($p < .0001$)

Table 4. Correlations for all associative pairs (n = 389).

Experiment 3

Given that there appears to be considerable semantic overlap between many of the associated word pairs used in the word "association" norms using in Experiment 2, it would be desirable to separate the semantic + associated pairs from the associated-only pairs. We approached this task by differentiating between word pairs which were "semantic neighbors" that can be calculated using HAL and those which were not.

To determine if a word is a semantic neighbor of another word, semantic distances are computed between the first word and all other words using HAL's semantic vectors. These distances are then ranked using this semantic distance. We arbitrarily chose to call the fifty words with the smallest distances "neighbors" of the target word, on the assumption that they are the most highly related words in the available vocabulary (which consists of the most common 70,000 tokens found in our corpus); see Table 5 for some examples of semantic neighborhoods.

Of the 389 original word pairs, 67 (~17%) qualified as semantic neighbors, leaving 322 as non-neighbors (although still all highly associated). As in Experiment 2, correlations were computed between the associativity ratings of these two sets of word pairs (neighbors and non-neighbors) and semantic distance, co-location separation, and co-location frequency.

woman	girl, man, child, piece, huge, woman's, cow
gallon	gallons, liter, inch, pound, megs, litre, meg
red	blue, green, white, black, gold, monster, ring
spider	turtle, shark, angel, dragon, storm, slug, giant

Table 5. Example semantic neighborhoods.

Results and Discussion

Correlations are shown in Table 6. Correlations between associativity and both semantic distance and co-location separation are similar, for both sets of words, to those obtained in Experiment 2. However, a striking difference is found in the correlations for co-location frequency. Here, the correlation for the semantic neighbor set has increased, by nearly a factor of two, to 0.48 ($p < 0.0001$), while the correlation for non-neighbors has become negligible and unreliable. This result suggests that the popular view that association is reflected by word proximity is only true for words which are semantically related (though, for these words, co-location frequency is an excellent predictor of associativity).

	N	Distance	Col. Sep.	Col. Freq.
Non-neighbors	322	0.03 (p = 0.552)	-0.08 (p = 0.16)	0.05 (p = 0.412)
Neighbors	67	-0.13 (p = 0.552)	-0.21 (p = 0.08)	0.48 (p < .0001)

Table 6. Correlations with associativity.

General Discussion

Experiments 2 and 3 validate the distinction between associated and associated-semantic pairs which was made in Experiment 1 (see Chiarello et al., 1990). Clearly, among associates produced by human subjects, there are both associates which co-occur in natural language and those which do not; the distinction appears to be quite sharp, with those not co-occurring in language being the substantial majority.

Temporal contiguity has been thought to be a critical component of learning (Deese, 1965). Since there was not a general correlation between associativity and co-location frequency, it seems unlikely that the majority of associations produced by subjects were learned via temporal contiguity in this fashion (at least from natural language); only those associations which also contain a semantic component appear to influence word order.

Experiment 1 demonstrated that, whatever the method is by which these non-temporal associations are expressed, HAL is not sensitive to them. An example in which HAL picks up semantic information which is not directly temporally expressed, is the relationship between *road* and *street*. The collocation frequency for this word pair is only 74 (as compared to 1497 for *man-woman*), yet the vector representations for *road* and *street* are nearly identical (indicating that they share a great deal of meaning).

The methodology and results presented here have practical applications; stimulus set construction, for instance, could benefit from an objective measure of semantic overlap in associative data. More importantly, this research sheds light on the nature of associativity, dispelling some popular ideas about where it originates. There are several important theoretical conclusions to be drawn from this set of experiments and the earlier work on HAL (Burgess & Lund, 1996; Lund & Burgess, in press; Lund et al., 1995). The reliance of our notions about semantic organization on associative structure is virtually axiomatic. George Miller (1969) proposed that "meanings can be characterized in terms of lexical associations. Lexical associations can be measured by word association tests. The results of word association tests can be accounted for in terms of particular types of sentences we can form with the words" (pp. 234-235). These beliefs have been persistent (Plaut, 1995; Shelton & Martin, 1992; Spence & Owens, 1990). The notion that associativity can be characterized by temporal association in language receives little or no support from our corpus analysis. The extent that "word associations" can be characterized by temporal association seems to be more a function of semantic neighborhood which is a reflection of semantic similarity in HAL's vector representations. This may be counterintuitive, given that HAL's methodology requires it to bootstrap its representations from lexical co-occurrence and that lexical co-occurrence is related to temporal association. The relationship between associative (and thus temporal) connections in memory and semantic representations corresponds to how a memory is initially acquired and how it is ultimately transformed into a more generalized piece of knowledge (its semantic vector in the HAL model). The notion that associations provide the antecedents for semantic structure was early posited in cognitive psychology (Mandler, 1962; Osgood, 1971). This is not to say that associative information may not be important in more generalized structures. For example, while there is not much similarity between the concepts *man* and *tall*, they are strongly associated and that men tend to be tall relative to some standard is important knowledge. However, associative information such as this tends to be more relativistic. An outcome of this is that priming effects using associative relationships may be less stable across sets of items or subjects, as we saw in experiment 1. Associative information that has not become generalized, that is, not semantic, would be more predictive of episodic relationships.

In the HAL model, this distinction between associative and semantic information corresponds to the distinction between *local* co-occurrence and *global* co-occurrence. Temporal information is reflected in local co-occurrence. When one examines global *patterns* of co-occurrence, across a vocabulary, one finds not associative, but semantic information. Only a small proportion of the vector elements are required to obtain the cognitive effects that we find. Thus, the functional representation is abstracted from and is more a measure of word contextuality or global co-occurrence than temporal association. This is an important distinction to make as theories of semantics develop (Burgess & Cottrell, 1995; Burgess & Lund, 1996; Landauer & Dumais, 1996; Lund & Burgess, in press).

Acknowledgments

This research was supported by NSF Presidential Faculty Fellow award SBR-9453406 to Curt Burgess. We thank Catherine Decker and three anonymous reviewers for their helpful comments.

References

Buchanan, L., Burgess, C., & Lund, K. (1996). *Over-crowding in Semantic Neighborhoods: Modeling Deep Dyslexia.* Paper submitted to TENNET VII, Seventh Annual Conference on Theoretical and Experimental Neuropsychology, Montreal, Quebec.

Burgess, C., & Cottrell, G. (1995). Using high-dimensional semantic spaces derived from large text corpora. *Proceedings of the Cognitive Science Society* (pp.13-14). Hillsdale, N.J.: Erlbaum Publishers.

Burgess, C., & Lund, K. (1996). *Parsing constraints and high-dimensional semantic space.* Manuscript submitted for publication.

Burgess, C., & Lund, K. (in press). Modeling cerebral asymmetries of semantic memory using high-dimensional semantic space. In Beeman, M., & Chiarello, C. (Eds.), *Getting it right: The cognitive neuroscience of right hemisphere language comprehension.* Hillsdale, N.J.: Erlbaum Press.

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 & Language, 38,* 75-104.

Collins, A.M., & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247.

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

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

Gernsbacher, M.A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: LEA.

Landauer, T.K., & Dumais, S. (1994). *Memory model reads encyclopedia, passes vocabulary test.* Paper presented at the Psychonomics Society.

Lund, K., & Burgess, C. (in press). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers.*

Lund, K., Burgess, C., and Atchley, R.A. (1995). Semantic and associative priming in high-dimensional semantic space. *Cognitive Science Proceedings,* 1995, 660-665

Lupker, S.J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning & Verbal Behavior, 23,* 709-733.

Mandler, G. (1962). From association to structure. *Psychological Review, 69,* 415-427.

McClelland, J.L., & Kawamoto, A.H. (1986). Mechanisms of sentence processing: assigning roles to constituents of sentences. *Parallel Distributed Processing, Volume 2.* Cambridge, MA.: MIT Press.

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology , 90,* 227-234.

Miller, G. (1969). The organization of lexical memory: Are word associations sufficient? In G. A. Talland & N. C. Waugh (Eds.), *The pathology of memory.* New York: Academic Press.

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.

Osgood, C. E. (1971). Exploration in semantic space: A personal diary. *Journal of social issues, 27,* 5-64.

Palermo, D.S., & Jenkins, J.J. (1964). *Word association norms grade school through college.* Minneapolis, MN: University of Minnesota Press.

Plaut, D.C. (1995). Semantic and associative priming in a distributed attractor network. *Cognitive Science Proceedings,* 1995, 37-42

Schvaneveldt, R.W. (1990). Pathfinder associative networks: Studies in knowledge organization. Norwood, NJ: Ablex.

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.