Distinguishing Between Manner of Motion and Inherently Directed Motion Verbs Using a High-dimensional Memory Space and Semantic Judgments

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

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

Abstract

Levin (1993) has proposed a semantic distinction between two types of motion verbs: manner of motion verbs and inherently directed motion verbs. In contrast, Jones (1995) has argued that this distinction is better accounted for by syntactic principles. Two simulations are presented that demonstrate that verb representations from the Hyperspace Analogue to Language (HAL) model of memory (Burgess & Lund, 1997a; Lund & Burgess, 1996) are sensitive to the distinction between these two verb classes. The second simulation shows that this effect is not due to word frequency differences. The final experiment uses human judgments of concreteness, imageability, and familiarity on these verbs to provide further data on the particular semantic characteristics that may be salient to the language user. We argue that these results provide empirical support for Levin's semantic distinctions.

Theories of how word meanings are represented in the mental lexicon must account for a range of phenomena. Among these is the fact that words that are similar in meaning tend to be used in similar ways in language. As well, the understanding (implicit or explicit) that particular words can be used in certain ways, necessarily, must be a characteristic of lexical knowledge (Levin, 1993). For example, while it is not uncommon to hear people speak of "trucking" something from one location to another (furniture perhaps), using the term "carring" in a similar context would be highly unusual. The example highlights that, whether or not language users are aware of it, the manner in which words are used in language reflects characteristics of word meaning and similarity that are a function of their contexts.

While the exact nature of lexical representations is unknown, various attempts have been made to determine the sorts of information that exist within the lexicon. Traditionally, lexical priming procedures, categorization experiments, and word association norms have been utilized to shed light on the nature of lexical knowledge (Neely, 1991; Collins & Loftus, 1975; Rips, Shoben, & Smith, 1975; Rosch, 1978). In addition, rigorous linguistic procedures have assembled a wealth of information about word meanings (Jackendoff, 1990; Miller & Johnson-Laird, 1976). More recently, researchers have applied computational and statistical techniques to large corpora of on-line text and machine-readable dictionaries to study lexical representations (Burgess & Lund, 1997a; Landauer & Dumais, 1996; Lund & Burgess, 1996; Miller, 1990).

One finding that seems to be consistent across a number of investigations is that lexical knowledge is organized into somewhat coherent "semantic fields," groups of words that share some degree of semantic, orthographic, phonological, or syntactic similarity (Levin, 1993; Miller & Johnson-Laird, 1976; Talmy, 1985). Accepting this notion greatly simplifies research of lexical concepts, as this allows the division of lexical space (i.e., all word representations) into smaller, more manageable units. If one assumes that one of the more consequential lexical divisions is based on grammatical class, then this provides the opportunity (and the motivation) to begin an investigation into nouns, or verbs, or adjectives, rather than having the range of all words as the only possible starting point.

Levin (1993) has recently undertaken the ambitious task of categorizing over 3,000 verbs into distinct classes based on verb semantics and syntactic behavior. Levin worked with the theoretical assumption that verb behavior is semantically determined (or, at the least, that verb syntax is predictably related to verb meaning), and that an intensive consideration of the similarity of verb meanings would enable the development of verb classes in which the members of a particular class would all share certain characteristics of meaning and behavior. By producing these classifications Levin sought to move closer to the eventual isolation of a complete set of "meaning components" that would bear upon the nature of lexical representations of verb meaning. To her credit, Levin produced 191 distinct classes, each of which captured a coherent collection of semantically similar verbs. Unfortunately, Levin's procedure for classification did not rest entirely upon verb meaning, as a great deal of consideration was given to the behavior of individual verbs with respect to various diathesis

alternations. Highlighting the distinction between verbs with regard to these alternations allowed Levin to exploit an important lexical similarity that is not wholly semantic in nature; and, thus, Levin's book (Part II) does not stand as the strongest test of the hypothesis that the meaning of a verb determines its syntactic behavior.

Jones (1995) evaluated the verb class distinctions proposed by Levin (1993) and concluded that classifying verbs based on semantic similarity was not the only way to produce such distinctions; and that, in fact, one could reproduce a large percentage of Levin's verb classes using a system based exclusively on verb syntax. Using the example sentences provided in Levin, Jones was able to derive "syntactic signatures" for individual classes based on the particular syntactic patterns that occurred within the sentences for a given class. When both positive (971 instances) and negative (554 instances) example sentences were included in deriving the syntactic signatures, Jones was able to "replicate" 80.1% of the verb class distinctions presented by Levin. This would be quite a compelling result if it were not for the fact that the inclusion of negative evidence is a cognitively unrealistic assumption when considering the human analogue to Jones' procedure--there is little evidence that the human language input stream has an abundance of "negative" sentences (Bowerman, 1987). Without the negative examples Jones was able to reproduce only 37.7% of Levin's classifications, a drop in precision that prompts one to conclude that the classification of semantically similar verbs based on syntactic patterns may be problematic.

The intuitions of trained analysts of lexical representations (i.e., linguists, lexicographers, psycholinguists) often stand as "educated guesses" as to the appropriateness of various distinctions between classes of words. However, though these divisions seem to have more than an air of correctness, such intuitions may be subject to theoretical and measurement bias. As such, these educated intuitions would be buttressed by more empirically rigorous procedures for testing the similarity or dissimilarity between putative classes of words. Objective methods for investigating word representations through statistical analysis of large-scale text corpora are presently being developed (Finch & Chater, 1992; Landauer & Dumais, 1996; Lund & Burgess, 1996). These procedures appear to have an obvious advantage over less technologically advanced techniques in that these measures have access to many hundreds of millions of words. More important than the immense size of these databases, however, is the fact that these procedures encounter written text much as it occurs in human language use. The representations which result come from an analysis of text alone. Expert theories do not produce these word representations, nor do these procedures rely on explicit linguistic "rules" of the sort human language users are aware. One anticipation is that, while present intuitions about the nature of word meanings and the relationship between meanings in the lexicon may not be wholly inaccurate, newer investigative methods may provide the means by which these intuitions can be subjected to empirical study, and if necessary, appropriate revision.

One of the more recent methods of representing word meanings is the Hyperspace Analogue to Language (HAL) model (Burgess & Lund, 1997a; Lund & Burgess, 1996). HAL stands as a credible attempt to produce word representations that encode various aspects of meaning. HAL has been used to account for grammatical class distinctions and semantic effects on syntactic processing (Burgess & Lund, 1997a), several semantic and associative priming effects (Lund & Burgess, 1996; Lund, Burgess, & Audet, 1996), the sort of semantic errors made by deep dyslexia patients (Buchanan, Burgess, Lund, 1996), and cerebral asymmetries in semantic memory processing (Burgess & Lund, 1997b).

HAL is a model that develops word meaning from global co-occurrence statistics extracted from human on-line language use. A ~300 million word corpus of Usenet text is the input stream from which HAL records weighted co-occurrence information for the 70,000 most frequent vocabulary items. The process of recording these co-occurrences allows for the formation of a co-occurrence matrix from which word vectors are derived. Conceptually, these vectors represent points in a high-dimensional space. Having developed high-dimensional representations for words, similarity between words corresponds inversely to inter-point distances. The assumption is that the more similar two words are, the closer their points in the high-dimensional space (see Burgess & Lund, 1997a, for further discussion of the HAL methodology).

An important consideration is the fact that HAL's representations are the product of a meaning acquisition process that encounters words as they occur in written language. Developing these meaning representations does not require any theory of semantic features, nor are these representations the products of expert linguistic evaluation. HAL's representations are a function of the global co-occurrence of lexical items in language use. This distinguishes HAL from those categorization schemes that are based on the explicit knowledge of word meanings and behavior held by lexical analysts. HAL does not "know the rules," as it were, nor does it encounter words in the rich environment in which we exist; it simply is sensitive to how words are used in relation to other words (i.e., context). This sensitivity allows HAL to develop word representations that may be used to test theoretically derived classification schemes that currently exist.

Of the numerous verb classes considered by Levin (1993), one of the larger classes, verbs of motion, was further subdivided by Levin into manner of motion verbs (e.g., *drift*, *slide*, *trot*) and inherently directed motion verbs (e.g., *arrive*, *depart*, *leave*). These verb classes were distinguished based on the notion that all of the manner of motion verbs have meanings that include some manner or means of

motion, whether the movement is by an animate (e.g., *trot, swim*) or inanimate (e.g., *twist, glide*) entity. Conversely, the meaning for each of the inherently directed motion verbs involves a specification of the direction of movement. This distinction between verb classes appears to be an appropriate test, not only of the ability of the HAL model to be sensitive to this dissociation, but perhaps a test of the appropriateness of the division between these two classes proposed by Levin. However, as Jones (1995) was able to reproduce a significant proportion of Levin's verb-class distinctions *only* when both positive and negative example sentences were included, one might predict that HAL will fail to distinguish between the two verb classes, as the HAL model rarely, if ever, encounters "negative" sentences.

What follows are two HAL simulations using the verbs Levin identified as either manner of motion verbs, or inherently directed motion verbs. In the first simulation, an attempt is made to reproduce Levin's distinction between the two verb classes with an alternate methodology and without relying on negative evidence as employed in Jones' approach. In the second simulation, the issue of frequency-sensitivity is introduced, and an attempt is made to produce the verb class distinction using a HAL matrix in which the relative frequencies of words have been controlled. Following the simulations, consideration is given to human judgements of word characteristics, and the role this information might play in distinguishing verb classes.

Experiment 1: Dissociating Manner of Motion Verbs from Inherently Directed Verbs

Dissociation of grammatical categories has been demonstrated before using HAL (Burgess & Lund, 1997a). The intent of the current simulation was to provide evidence that the HAL model is sensitive to finer-grained distinctions between members of a single grammatical class, namely motion verbs. Two classes of motion verbs--manner of motion verbs and inherently directed motion verbs--were subjected to multidimensional scaling (MDS) using HAL's word vectors as a set of coordinates in a high-dimensional space in order to show that the interword distances in the space can provide a basis for distinguishing between these verb classes.

Method

Materials. Levin (1993) presented 22 verbs of inherently directed motion, while 138 manner of motion verbs were listed. Sixty of the manner of motion verbs did not occur in the HAL matrix; from the remaining 78 verbs, 20 were randomly chosen for the simulation. Of the 22 inherently directed verbs, 12 were ultimately chosen for the simulation; six verbs did not occur in the matrix. *Tumble* was removed due to low frequency in the HAL corpus, and *cross* was removed because Levin explained that this verb does not behave like the other verbs in the class. *Exit* and *advance* were removed due to the author's intuition that these words might behave erratically in the simulations due to their high frequency use as nouns (pilot simulations confirmed that *advance* and *exit* do cluster erratically in the MDS solution with this set of items).

Procedure. Co-occurrence vectors were extracted from the HAL model for the verbs. Each vector was treated as a set of coordinates in a high-dimensional Euclidean space, and a MDS solution was computed. The hypothesis was that these word vectors, representing the interword distances for the chosen set of words, would operate as a similarity matrix (Lund & Burgess, 1996).

Results

The similarity matrix was analyzed by a MDS algorithm that projects points from a high-dimensional space into a lower-dimensional space in a nonlinear fashion that attempts to preserve the distances between points. The lower-dimensional projection allows for the visualization of the spatial relationships between the co-occurrence vectors. The two-dimensional MDS solution for the verbs is shown in Figure 1.

Visual inspection of Figure 1 suggests that the two verb classes were differentiated in the MDS solution. The manner of motion verbs appear to occupy a separate space than the

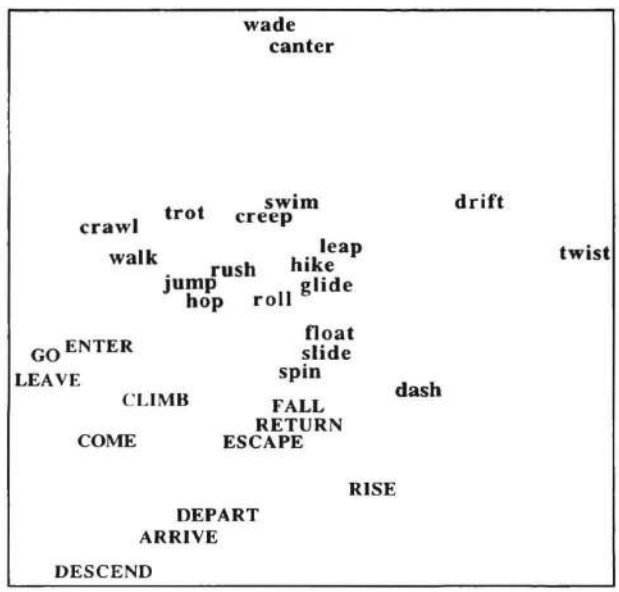

Figure 1: Two-dimensional multidimensional scaling solution for word vectors for manner of motion verbs (lowercase) and inherently directed motion verbs.

verbs of inherently directed motion. However, given the nature of the MDS procedure, in which an extreme reduction of dimensionality occurs when projecting data from a high-dimensional space down to only two dimensions, it was necessary to perform appropriate inferential statistics. An

analysis of variance was performed which compared intragroup distances between verbs with intergroup distances between these items. Distances between all combinations of verb pairs within a group (i.e., verb class) were calculated and compared to the distances between all combinations of verb pairs between groups. Manner of motion verbs were differentiated from inherently directed motion verbs, $F(1, 618) = 7.37$, $p = .0068$. As well, verbs of inherently directed motion were differentiated from manner of motion verbs, $F(1, 370) = 72.94$, $p < .0001$.

Discussion

Verb-class information carried in HAL's vector representations is sufficient to distinguish between two groups of motion verbs that have some degree of semantic dissimilarity, as proposed by Levin (1993). These results support Levin's theoretical distinction between manner of motion verbs and inherently directed motion verbs.

Experiment 2: Evaluating Frequency Effects in the Dissociation between Manner of Motion Verbs from Inherently Directed Verbs

The claim has been made that all models of cognitive processes that are sensitive to the statistical properties of the system's input are frequency-sensitive rather than reflecting the richer properties of language use (Fodor & Pylyshyn, 1988). As the HAL model relies upon the statistical aggregation of lexical co-occurrences, it falls squarely within the range of models that potentially owe a great debt to the relative frequencies of lexical items occurring in the input stream. Therefore, it is important to evaluate the degree to which the relative frequencies of lexical items produce the effects shown with the HAL model. Inherently directed motion verbs were more frequent (259 occurrences per million words) than manner of motion verbs (31 occurrences per million), ($t(30) = 2.49$, $p = .018$), using the original HAL corpus of 131 million words (Burgess & Livesay, in press). This difference motivates the efforts of Experiment 2 in which the simulation procedure will be repeated using a matrix that has been controlled for differences in word frequency.

Method

Materials. Stimuli used in Experiment 2 were identical to that used in Experiment 1.

Procedure. A new matrix was formed by collecting co-occurrence counts for the 32 stimulus items. To control for frequency, co-occurrence statistics were collected for each word until that item had a raw frequency count in the matrix of 800 occurrences (854 was the frequency for the lowest frequency verb). All subsequent procedures involving the development of a similarity matrix for the 32 stimulus items were identical to those performed in Experiment 1.

Results

The two-dimensional MDS solution for the 32 verbs is shown in Figure 2. The two-dimensional separation of the two verb classes was not as distinct as seen in Experiment 1, however, the majority of inherently directed motion verbs clustered together, as they had in the previous MDS.

As in Experiment 1, an evaluation of the high-dimensional separation of the verb classes required a comparison of the intragroup distances between verbs with the intergroup distances between these items. Verbs of inherently directed motion were differentiated from manner of motion verbs, $F(1, 370) = 54.95$, $p < .0001$, however, manner of motion verbs were not reliably differentiated from inherently directed motion verbs due to the migration of three verbs to the other verb class space, $F(1, 618) = 1.29$, $p = .25$.

Discussion

The notion that the relative frequency of lexical items in the input stream plays an important role in cognitive models such as HAL received some support in the current

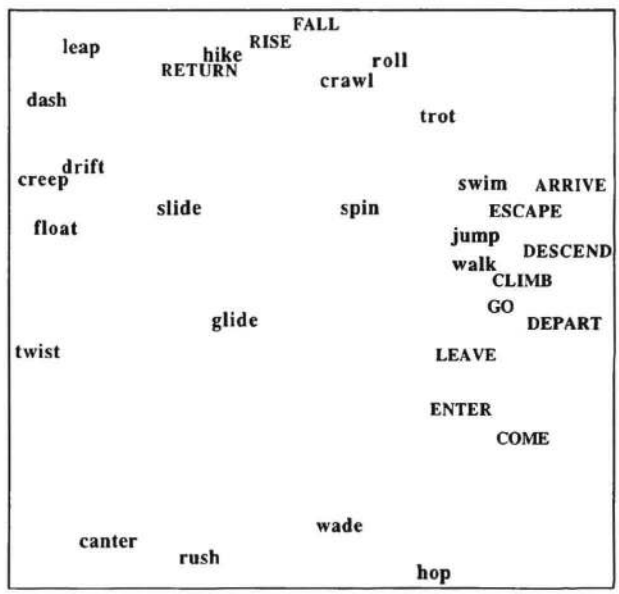

Figure 2: Two-dimensional multidimensional scaling solution for word vectors (for manner of motion verbs and inherently directed motion verbs) derived from a frequency-controlled matrix.

simulation. When HAL encountered every stimulus item exactly 800 times the resulting word representations did not possess the robust contextual experience required to completely dissociate inherently directed motion verbs from manner of motion verbs. However, in spite of the restricted nature of these word representations, the contextual information that was present was sufficient to allow the majority of inherently directed motion verbs to cluster

together. Limiting the experience HAL has with individual words by restricting word frequency has the consequential effect of also limiting the range of contexts in which the model encounters these words. Thus, it is not surprising that the verb-class dissociation seen in Experiment 1 was not quite as pronounced in the current simulation.

Experiment 3: Distinguishing Motion Verbs by Human Semantic Judgments

The results of Experiments 1 and 2 independently verify the distinction between the two motion verb classes made by Levin (1993). However, what is unclear is exactly what characteristics of these verbs serve to distinguish the two classes. Levin proposed the distinction between the groups of verbs based on an important semantic dissimilarity (e.g., characteristics of motion). Using the HAL model, the word vectors for these verbs provide a set of co-ordinates in high-dimensional meaning space that demonstrates that these two verb types indeed exist in separate categorical spaces as suggested by Levin. In HAL, the vector elements that are the co-ordinates in the meaning space are other symbols in the original input stream. To further explore what semantic differences there may be between these two classes of motion verbs, we compared the two sets of verbs using Toglia and Battig's (1978) *Handbook of Semantic Word Norms*. These norms provide a database of human judgments about words on such semantic indices as concreteness, imageability, and meaningfulness. The two classes of verbs were compared on these semantic indices to determine what characteristics might discriminate between them.

Methods

Human ratings were extracted from the Toglia and Battig norms for the 14 verbs from Experiment 1 that occurred in the Toglia and Battig database (e.g., *come, crawl, fall, float, go, jump, leap, rise, roll, rush, slide, swim, twist, walk*). These norms included ratings on: concreteness, imageability, categorizability, meaningfulness, familiarity, number of attributes, and pleasantness.

Results

Manner of motion verbs were rated higher than verbs of inherently directed motion on concreteness ($t(12) = 2.75$, $p = .018$), imageability ($t(12) = 2.21$, $p = .047$), and categorizability ($t(12) = 2.51$, $p = .028$). In contrast, inherently directed verbs were rated more highly on familiarity ($t(12) = 2.53$, $p = .027$). The two verb classes did not differ on meaningfulness ($t(12) = .15$, $p = .88$), number of attributes ($t(12) = 1.41$, $p = .19$), or pleasantness ($t(12) = .1$, $p = .92$). These results indicate that, at least for the verbs that did occur in the Toglia and Battig norms, these two classes of verbs differ on four semantic characteristics.

Discussion

Introspection of the two verb classes suggests that the manner of motion verbs are more easily visualized than the inherently directed motion verbs. Visualizing a person swimming, spinning, or crawling intuitively appears to involve a more straightforward context than does arriving, entering, or rising. This intuition is supported by the normative results that the manner of motion verbs were rated as more concrete and imageable by human raters. On the other hand, the verbs of inherently directed motion were rated as more familiar. The exact relationship between the normative information and human language use is not completely clear, particularly since only 14 of the 32 motion verbs occurred in the Toglia and Battig norms (and only four of the inherently directed motion verbs). However, these apparent differences in human judgments between the two verb classes is supported by the distinction between the two verb groups in the high-dimensional analysis (see Figure 1). In an earlier experiment, Burgess and Lund (1997a) used HAL's word vectors in a multiple regression procedure to predict the human ratings for concreteness and imageability. A substantial amount of variance was accounted for with the word vectors for concreteness ($R^2 = .65$) and imageability ($R^2 = .63$)--both factors upon which these two classes of motion verbs differ. We suspect that the results from Experiment 3 point to several factors that may be salient in discriminating between these two classes of motion verbs whether by computational approaches (see Figure 1) or by linguistic judgement (Levin, 1993).

General Discussion

Previous research has demonstrated that HAL's representations are sensitive to the sorts of meaning information that serve to distinguish words of different grammatical class (Burgess & Lund, 1997a). In addition to distinguishing between nouns, verbs, prepositions, and determiners, the HAL model was also able to distinguish between different morphological forms of verbs. Unambiguous past-tense verbs (e.g., *grew*) were distinguished from unambiguous past-participle verbs (e.g., *shown*). In the current study a group of motion verbs was subdivided into two meaningfully distinct classes of verbs: verbs that concern the manner in which an entity moves, and verbs which entail a particular direction of a motion without further specification by prepositions phrases (i.e., *away from*, *up onto*, etc.). These results add to the range of phenomena concerning word meanings to which the HAL model is responsive.

Contrary to Jones' (1995) suggestion that negative examples play a critical role in dissociating the two verb classes, the HAL model encounters few, if any, "negative sentences" of the sort Jones included in his evaluation of Levin's verb classes. These findings provide further support for the notion that using a global co-occurrence procedure such as HAL's that produces word representations based

solely on the system's experience with incoming text is a robust way in which to derive word meanings. The results of Experiment 3 suggest four semantic variables that may play a role in a person's ability to distinguish between these two verb classes: concreteness, imageability, categorizability, and familiarity.

The task of isolating distinct classes of verbs that share certain aspects of meaning may provide some evidence for the linguistically meaningful set of components which determine the lexical representations for these items in memory. Each class of verbs entails a set of lexical properties which reflects the sum total of meaning components across the members of a particular class. The aggregate properties displayed by a class may correspond to the types of lexical information available in word representations in memory, as these shared properties appear to have a great deal to do with how particular verbs are used in language (Levin, 1993). We conclude that (1) degrees of contextual similarity between verbs serve to allow intuitive separation of these motion verbs into meaningful groups, and (2) that differences in meaning between groups are systematically related to how verbs from different classes are used differently in language (Burgess & Lund, 1997a; Levin, 1993; Levin & Rappaport Hovav, 1991).

Acknowledgements

This research was supported by NSF Presidential Faculty Fellow award SBR-9453406 to Curt Burgess. We thank Michael Toglia and William Battig for supplying a machine readable version of their Handbook used in Experiment 3.

References

Buchanan, L., Burgess, C., & Lund, K. (1996). Overcrowding in semantic neighborhoods: Modeling deep dyslexia. *Brain and Cognition, 32*, 111-114.

Burgess, C., & Livesay, K. (in press). The effect of corpus size in predicting reaction time in a basic word recognition task: Moving on from Kucera and Francis. *Behavior Research Methods, Instruments, and Computers.*

Burgess, C., & Lund, K. (1997a). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes, 12*, 1-34.

Burgess, C., & Lund, K. (1997b). Modeling cerebral asymmetries in high-dimensional semantic space. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-28.

Finch, S., & Chater, N. (1992). Bootstrapping syntactic categories by unsupervised learning. *Proceedings of the Cognitive Science Society* (pp. 820-825). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jackendoff, R. S. (1990). *Semantic structures.* Cambridge, MA: MIT Press.

Jones, D. (1995). *Predicting semantics from syntactic cues--An evaluation of Levin's English verb classes and alternations*. (Tech. Rep. No. UMIACS-TR-95-121). University of Maryland, Institute for Advanced Computer Studies.

Landauer, T. K. & Dumais, S. T. (1996). How come you know so much? From practical problem to theory. In D. Hermann, C. McEvoy, M. Johnson, & P. Hertel (Eds.), *Basic and applied memory: Memory in context* (pp. 105-126). Mahwah, NJ: Erlbaum

Levin, B. (1993). *English verb classes and alternations.* Chicago: University of Chicago Press.

Levin, B. & Rappaport Hovav, M. (1991). Wiping the slate clean: A lexical semantic exploration. *Cognition, 41*, 123-151.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instrumentation, and Computers, 28*, 203-208

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

Miller, G. A. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography, 3*, 235-312.

Miller, G. A. & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Neely, J. (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* (pp. 264-336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1-20.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. Lloyd (Eds.) *Categorization and cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Talmy, L. (1985). Lexicalization patterns: Semantic structures in lexical forms. In T. Shoben (Ed.) *Language typology and Syntactic description, 3, Grammatical categories and the lexicon* (pp. 57-149). New York: Cambridge University Press.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Lawrence Erlbaum Associates.