To appear in In A. Healy (Ed.), Experimental Cognitive Psychology and its Applications: Festschrift in Honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer.

Word Association Spaces for Predicting Semantic Similarity Effects in Episodic Memory

Mark Steyvers* **Richard M. Shiffrin** **Douglas L. Nelson**

University of California at Irvine Indiana University University of South Florida

A common assumption of theories of memory is that the meaning of a word can be represented by a vector which places a word as a point in a multidimensional semantic space (e.g. Landauer & Dumais, 1997; Burgess & Lund, 2000; Osgood, Suci, & Tannenbaum, 1957). Representing words as vectors in a multidimensional space allows simple geometric operations such as the Euclidian distance or the angle between the vectors to compute the semantic (dis)similarity between arbitrary pairs or groups of words. This representation makes it possible to make predictions about performance in psychological tasks where the semantic distance between pairs or groups of words is assumed to play a role.

One recent framework for placing words in a multidimensional space is Latent Semantic Analysis or LSA (Derweester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). The main assumption is that the similarity between words can be inferred by analyzing the statistical regularities between words and text samples in which they occur. For example, a textbook with a paragraph that mentions “cats” might also mention “dogs”, “fur”, “pets” etc. This knowledge can be used to infer that “cats” and “dogs” are related in meaning. The technique underlying LSA is singular value decomposition (SVD). This procedure is applied to the matrix of word-context frequencies in a high dimensional space (typically with 200-400 dimensions) in which words that appear in similar contexts are placed in similar regions of the space. Interestingly, some words that never occur in the same context might still be similar in LSA space if they co-occurred with other words that do occur together in text samples. Landauer and Dumais (1997) applied the LSA approach to over 60,000 words appearing in over 30,000 contexts of a large encyclopedia. More recently, LSA was applied to over 90,000 words appearing in over 37,000 contexts of reading material that an English reader might be exposed to from 3rd grade up to 1st year of college from various sources such as textbooks, novels, and newspaper articles. The LSA representation has been successfully applied to multiple choice vocabulary tests, domain knowledge tests and content evaluation (see Landauer & Dumais, 1997; Landauer et al. 1998).

In this research, we will apply scaling techniques such as SVD as well as Multidimensional Scaling on a large database of free association collected by Nelson, McEvoy, and Schreiber (1999) containing norms for first associates for over 5000 words. By applying scaling methods on the free association norms, we hope to uncover the latent information available in the free association norms that is not directly available by investigating simple measures for associative strengths based on the direct and indirect associative strengths through short chains of associates (e.g., Nelson & Zhang, 2000). The basic approach is illustrated in Figure 1. The free association norms were represented in matrix form with the rows representing the cues and the columns representing the responses. The entries in the matrix are filled by some measure of associative strength between cues and responses. By applying scaling methods on the matrix, words are placed in a high dimensional space such that words with similar associative patterns are placed in similar regions of the space. We will refer to the resulting space as the

*Send correspondence to: Mark Steyvers, Department of Cognitive Sciences, 3151 Social Sciences Plaza, University of California, Irvine, CA 92697-5100. msteyver@uci.edu

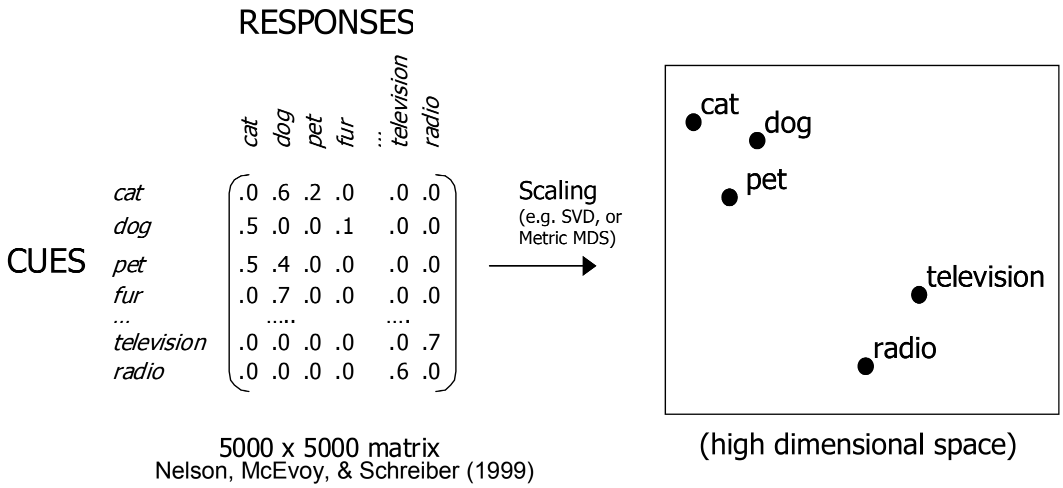

Figure 1. Illustration of the creation of Word Association Spaces (WAS). By scaling the word associations of a large database of free association norms, words are placed in a high dimensional semantic space. Words with similar associative relationships are placed in similar regions of the space.

Word Association Space (WAS).

We believe such a construct will be very useful in the modeling of episodic memory phenomena because the associative structure of words plays a central role in recall (e.g. Bousfield, 1953; Deese, 1959a,b, 1965; Jenkins, Mink, & Russell, 1958), cued recall (e.g. Nelson, Schreiber, & McEvoy, 1992), and recognition (e.g. Nelson, Zhang, & McKinney, 2001). For example, Deese (1959a,b) found that the inter-item associative strength of the words in a study list can predict the number of words recalled, the number of intrusions, and the frequency with which certain words intrude. In the present research, we will compare the performance of LSA with WAS in three episodic memory tasks: recognition memory, free recall and cued recall. It was expected that the similarity structure in WAS is well suited to predict various semantic similarity effects in these episodic memory tasks.

Word Association Spaces

Deese (1965) asserted that free associations are not the result of haphazard processes and that they arise from an underlying regularity in pre-existing associative connections. He laid the framework for studying the meaning of linguistic forms that can be derived by analyzing the correspondences between distributions of responses to free association stimuli: "The most important property of associations is their structure - their patterns of intercorrelations" (Deese, 1965, p.1). Deese applied factor analyses to the overlap in the distribution of free association responses for a small set of words and argued that these analyses could be used to learn about the mental representation of words. In this paper, we capitalized on Deese’s ideas of utilizing the pattern of intercorrelations in the free association norms by placing a large number of word associations in a semantic space and then used them to predict semantic similarity effects in memory. Instead of factor analyses, we used the techniques of singular value decomposition (SVD) and metric multidimensional scaling analyses.

The data for these procedures relied on free association norms involving more than 5,000 words and 6,000 participants (Nelson et al., 1999). An average of 149 (SD = 15) participants were each presented with 100-120 English words. These words served as cues (e.g. “cat”) for which participants had to write down the first word that came to mind (e.g. “dog”). For each cue the proportion of subjects that elicited the response to the cue was calculated (e.g. 60% responded with “dog”, 15% with “pet”, 10% with “tiger”, etc).

Scaling by Singular Value Decomposition

The method of SVD can be applied to any matrix containing some measure of strength or co-occurrence between two words. Although many different ways have been proposed to calculate an index of associative strength between two words (e.g., Marshall & Cofer, 1963; Nelson & Zhang, 2000), we will restrict ourselves to two simple measures of associative strength. Let A_{ij} represent the proportion of subjects that gave the response *j* to the cue *i*. The simplest measure would be to take A_{ij} itself. In the norms, the associative strengths A_{ij} are often highly asymmetric where the associative strength in one direction is strong while it is weak or zero in the other direction. Even though SVD can be easily applied to asymmetric matrices, the results are more interpretable when it is applied to symmetric matrices. Therefore, in our first measure for associative strength we take:

$$S_{ij}^{(1)} = A_{ij} + A_{ji}$$

$S^{(1)}_{ij}$ is equivalent to adding forward strength to backward strength. This measure is of course symmetric so that $S^{(1)}_{ij} = S^{(1)}_{ji}$. This measure is based on only the direct association between *i* and *j* and involves only one associative step going from *i* to *j* (hence the index '1'). In the norms of Nelson et al. (1998), subjects were only allowed to give the first response that came to mind. The second strongest response in one subjects' mind might be elicited by another subject or it might not be elicited at all if the first response is a strong associate. Therefore, the $S^{(1)}$ measure might be underestimating the associative strength between two words especially in cases where the measure is zero (Nelson et al., 1998). In the second measure for associative strength, we take:

$$S_{ij}^{(2)} = S_{ij}^{(1)} + \sum_k S_{ik}^{(1)} S_{kj}^{(1)}$$

This equals the forward plus backward plus mediated strength through other associates. Note that this measure involves the direct strength between i and j as well as the indirect strength by summing over all paths from *i* to *k* to *j*, the product of the symmetric associative strengths between *i* and *k*, and *k* and *j*. These indirect associative strengths involve the two step probabilities of going from *i* to *j* and vice versa (hence the index '2'). Research has shown that the indirect associative strengths play a role in cued recall (Nelson & Zhang, 2000) and recognition (Nelson, Zhang, & McKinney, 2001). For example, Nelson & Zhang (2000) found that including the indirect associative strengths in a measure for associative strength significantly increases the explained variance in the extra-list cued recall task.

We applied SVD separately on these two measures of associative strength. The result of each SVD is the placement of words in a high dimensional space, so that words that have similar associative structures are represented by similar vectors. Because of the SVD method, and based on work in LSA (see Derweester et al., 1990), a suitable measure for the similarity between two words is the cosine of the angle between two word vectors. Let $\bar{X}_i$ represent the vector in WAS for word *i*. The similarity between words *i* and *j* is calculated by:

$$similarity(i, j) = \cos(\alpha) = \frac{\bar{X}_i \cdot \bar{X}_j}{\left\|\bar{X}_i\right\| \left\|\bar{X}_j\right\|}$$

where $\left\|\bar{X}\right\|$ is the length of the vector and $\bar{X}_i \cdot \bar{X}_j$ represents the inner product between vectors *i* and *j*. Two words that are similar in meaning or that have similar associative structures are expected to have high similarity as defined by the cosine of the angle between the two word vectors. The SVD of the associative strengths can uncover the latent relationships between words. In the SVD of $S^{(1)}$, words that are not direct associates of each other can still be represented by similar vectors if their associates are related. In the SVD of $S^{(2)}$, words that not directly associated or indirectly associated through one intermediate associate, can still be represented by similar vectors if the associates of the associates of the words are related. In other words, the whole pattern of direct and indirect correlations between associations is taken into account when placing words in the semantic space.

An important variable is the dimensionality of the space. One can think of the dimensionality as the number of feature values for the words. The number of dimensions, which we varied between 10 and 500 will determine how much the information of the free association database is compressed. With too few dimensions, the similarity structure of the resulting vectors does not capture enough detail of the original associative structure in the database. With too many dimensions or the number of dimensions approaching the number of cues, the information in the norms is not compressed enough so that we might expect that the similarity structure of the vectors does not capture enough of the indirect relationships in the associations between words. In the analyses of predicting performance in a variety of tasks (recognition, free and cued recall), we will show that although the optimal number of dimensions depends on the specific task, intermediate values between 200 and 500 are appropriate for this method.

Scaling by Metric-MDS

An interesting comparison for the two WAS spaces based on SVD would be to construct a metric space in which the distance between two words, i.e., their dissimilarity, can be measured by the Euclidian distance between their vectors. Metric MDS is a classic method for placing stimuli in a space such that the Euclidian distance between points in the space approximates the Euclidian distances in the dissimilarity matrix. In order to apply metric MDS, estimates are needed for the distance between any two words. In fact, all non-diagonal entries in the matrix have to be filled with some estimate for the distance between words since no missing values are allowed in the method. This raises the problem how to estimate the distance between i and j when the associative strength as measured by $S_{ij}^{(1)}$ is zero.

In our solution of this problem, we were inspired by network models for proximity data (e.g. Cooke, Durso, & Schvaneveldt, 1986; Klauer, & Carroll, 1995). In these network models, dissimilarity between two stimuli is calculated by the shortest path between two nodes in a graph. In this research, we can use the word association norms as defining a graph: two words are linked by an edge if they have nonzero associative strengths. We will use the symmetric $S^{(1)}$ associative strengths because in the graph defined by $S^{(1)}$, it is possible to reach any word from any other word in the graph (in fact, the maximum number of steps between any pair of words is four). The distance between two words will be defined as the negative logarithm of the product of the associative strengths along the shortest path in the network defined by $S^{(1)}$. This is equivalent to the (negative) sum of the logs of the associative strengths along the shortest path:

$$T_{ij} = -\log\left(S_{ik}^{(1)} S_{kl}^{(1)} \cdots S_{qj}^{(1)}\right) = -\left[\log S_{ik}^{(1)} + \log S_{kl}^{(1)} + \ldots + \log S_{qj}^{(1)}\right]$$

Here, the shortest path between words i and j is from i to k to l through other words to q and finally j. With this distance measure, word pairs with weak or long associative paths are assigned large distances whereas word pairs with short or strong associative paths are assigned small distances. The distances T_{ij} were calculated for all word pairs in the word association database. Then, these distances were scaled by metric-MDS. The result is that the words are placed in a multidimensional space and the dissimilarity or distance between two words is expressed by the Euclidian distance between the two corresponding word vectors:

$$distance(i,j) = \left[\sum_k \left(X_{ik} - X_{jk}\right)^2\right]^{1/2}$$

Because of computational constraints, it was not possible to apply metric-MDS to the full matrix T containing the distances for all word pairs. Instead, we chose 2500 words from the original 5018 words in the word association database. The words in this smaller set included words appearing in various experiments listed in the next section and included a selection of randomly chosen words from the original set. As with the SVD scaling procedure, the number of dimensions was varied between 10 and 500.

Predicting Semantic Similarity Effects in Memory

Since Deese's (1959b) classic study on intrusions in free recall, many studies have shown that memory errors are in part based on semantic overlap between the response and the contents of memory. We introduced WAS as a way of quantifying the semantic similarity between words that might help in predicting these memory errors. Three sets of data were taken to assess the performance of WAS: a recognition memory experiment, Deese's original free recall experiment and a cued recall experiment. We tested three WAS based measures for semantic similarity. The first two were based on the SVD of $S^{(1)}$, the one step symmetric associative strengths, and on the SVD of $S^{(2)}$, the one plus the two step associative strengths involving indirect associative strengths. In these two semantic spaces (as in LSA) the cosine of the angle between two words expresses the similarity between two words. The last WAS measure was based on metric-MDS of the shortest path associative strengths. In this space, the Euclidian distance between two word vectors is taken as a measure for the dissimilarity between two words. These WAS scaling solutions were contrasted with the (unscaled) associative strengths $S^{(1)}$ and $S^{(2)}$ that were taken as control comparisons. We also tested two LSA based measures, one was based on a corpus of an encyclopedia and another on a corpus called tasa that included reading material that an English reader might be exposed to from 3rd grade up to 1st year of college.

Recognition Memory: Semantic Similarity Ratings

In an unpublished study by the first two authors (Steyvers & Shiffrin, Experiment 1), 89 subjects studied 144 words that contained 18 semantic categories of 5 words each. Based on a study by Brainerd and Reyna (1998), subjects gave two ratings for each of 100 test items. In one rating, they were instructed to judge whether the item was old or new and were told to judge semantically similar distractors as "new". In another rating, they were instructed to rate (on a six point scale) how semantically similar the item was to the studied items. We focused on the semantic similarity ratings for the new items from this study. For each subject, the 72 new test items were randomly selected from a larger pool of 144 words. An average of 44 (SD=4.87) subjects rated the semantic similarity for each of the 144 words that might appear as new words in the test list. The semantic similarity ratings are theoretically interesting because they can be used to test models of semantic similarity. Subjects merely have to remember how similar the item was to the studied items without being forced to give old-new judgments that might be more influenced by various

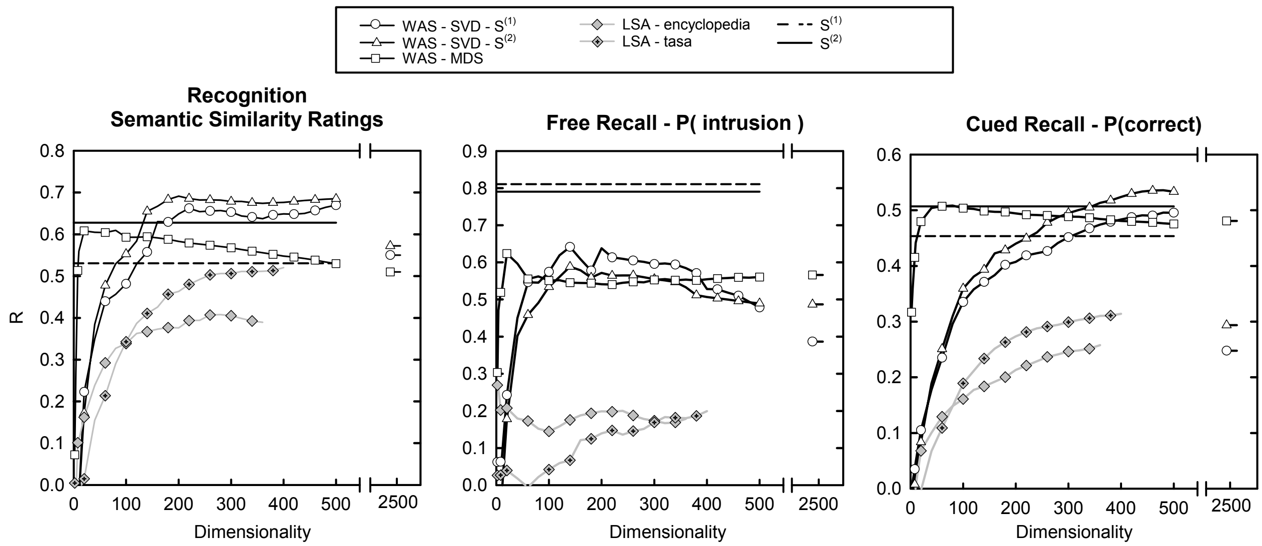

Figure 2. Correlations of different measures of semantic similarity for different dimensionalities. Data are taken from recognition memory, cued recall, free recall. See text for details.

strategic retrieval factors (such as word frequency or previous retrievals).

Many memory models assume that a recognition memory judgment is produced by calculating the global familiarity involving the summed similarity between the test item and the episodic traces in memory (e.g. Hintzman 1988; Gillund & Shiffrin, 1984). More recently, Shiffrin and Steyvers (1997, 1998) and McClelland & Chappell (1998) have proposed recognition memory models that produce recognition judgments with Bayesian decision processes. McClelland & Chappell (1998) proposed that the best match (i.e., maximum similarity) between the test item and the episodic traces in memory forms the basis for the recognition judgment. Shiffrin & Steyvers (1998) showed that in the Bayesian framework, a maximum similarity process produced results very similar to a summed similarity process. In this research, our aim is not to test these models specifically but to use and simplify the underlying mechanisms to predict semantic similarity ratings.

Inspired by the global familiarity and Bayesian recognition memory models, we measured the correlations between the semantic similarity ratings in the recognition memory experiment with the *sum* or *maximum* of the WAS similarity between the test item and all study words. Because the results were very similar for the sum and maximum calculations, we will list only the results for the maximum calculation.

The top left panel of Figure 2 shows the correlations between maximum similarity and number of dimensions (10-500) for the three WAS and two LSA based measures. For the SVD based semantic spaces, increasing the number of dimensions in either WAS or LSA increases the correlation generally up to around 200-300 dimensions. For WAS, an additional data point was plotted for 2500 dimensions which is the maximum number of dimensions given that the matrix contained only 2500 words (see previous section). This data point for 2500 dimensions was included because it represents the case where none of the indirect relationships in word association matrix are exploited and as such, no dimensionality reduction is performed. As can be observed, the correlation is lower for 2500 dimension indicating that some dimensionality reduction is needed to predict the semantic similarity ratings. Also, the SVD based on $S^{(2)}$ led to better correlations than the SVD based on $S^{(1)}$. This implies that adding the indirect associations in a measure for associative strength helps in predicting recognition memory performance. The two horizontal lines in the plot indicate the correlation when the associative strengths $S^{(1)}$ and $S^{(2)}$ are used as a measure for semantic similarity. The correlation is higher for $S^{(2)}$ than $S^{(1)}$ which again implies that in recognition memory, the indirect associative strengths help in predicting performance. Interestingly, the SVD scaling of $S^{(2)}$ gave higher correlations than associative strengths $S^{(2)}$ themselves. Even though $S^{(2)}$ includes the forward, backward and all two step associative strengths, applying the SVD and reducing the redundancies in the matrix of $S^{(2)}$ helped to increase the correlation. In other words, the indirect relationships and patterns of correlations that go beyond those of the two step associative strengths were utilized by the SVD procedure and these were beneficial in predicting the ratings from this recognition memory experiment.

The metric-MDS solution shows quite a different pattern of results than the SVD solution. The best correlation was obtained with 20-40 dimensions which is much lower than the number of dimensions typically needed in the SVD solutions of either WAS or LSA. Although the best correlation for metric-MDS was 0.6 as opposed to 0.7 for the SVD based solutions, it is interesting that relatively good performance can be achieved in semantic spaces that are of low dimensionality. Although specifying why this effect occurs is outside the scope of this paper, it could be related to the estimates involving the shortest associative path between words. As described in the previous section, in order to apply metric-MDS, estimates were needed for the distances between *all* word pairs in the vocabulary. The shortest associative path distance was proposed to meet this requirement; estimates were even generated for word pairs that were not associated directly or even indirectly through a chain of two associates. In SVD, no such estimates are required and those entries were left at zero. It is possible then, that the filling in process of all word pair dissimilarities by the shortest associative path distances helped in the global placement of all words in the semantic space.

Of the two corpora in LSA, the tasa corpus led to much better performance than the encyclopedia corpus. This difference is not surprising since the tasa corpus includes material that reflects much more closely the reading material an English reader is exposed to which in turn might lead to semantic spaces that are more psychologically plausible in terms of predicting semantic similarity effects in recognition memory. Comparing WAS to LSA, it becomes clear that WAS leads to much higher correlations than LSA.

Predicting Extralist Cued Recall

In extra-list cued recall experiments, after studying a list of words, subjects are presented with cues that can be used to retrieve words from the study list. The cues themselves are novel words that were not presented during study, and typically each word is associatively and/or semantically related to one of the studied words. The degree to which a cue is successful in retrieving a particular target word is a measure of interest because this might be related to the associative/semantic overlap between cues and their targets. Research in this paradigm (e.g., Nelson, Schreiber, & McEvoy, 1992; Nelson, McKinney, Gee, & Janczura, 1998; Nelson & Zhang, 2000) has shown that the associative strength between cue and target is one important predictor for the percentage of correctly recalled targets. Therefore, we expect that the WAS similarity between cues and targets are correlated with the percentages of correct recall in these experiments. We used a database containing the percentages of correct recall for 1115 cue-target pairs from over 29 extralist cued recall experiments from Doug Nelson's laboratory (Nelson, 2000; Nelson & Zhang, 2000).

The correlations between the various measures for semantic similarity and the observed percentage correct recall rates are shown in the rightmost panel of Figure 2. Overall, the results are very similar to the results obtained for the recognition memory experiment. The WAS space based on $S^{(2)}$ led to better performance than the WAS space based on $S^{(1)}$. Also, the associative strengths $S^{(2)}$ leads to better performance then the $S^{(1)}$ associative strengths. These findings are consistent with findings by Nelson & Zhang (2000) that show that the indirect relationships in word association norms can help in predicting cued recall performance. Interestingly, the plot also shows that the WAS space based on $S^{(2)}$ does somewhat better than the associative strengths $S^{(2)}$ it was based on. This advantage implies that applying dimensionality reduction to make greater use of the indirect associative connections helped in predicting cued recall. Finally, as with the recognition results, the WAS space correlates better with cued recall than LSA.

Predicting Intrusion Rates in Free Recall

In a classic study by Deese (1959b), the goal was to predict the intrusion rates of words in free recall. Fifty participants studied the 12 strongest associates to each of 36 critical lures while the critical lures themselves were not studied. In a free recall test, some critical lures (e.g. "sleep") were falsely recalled about 40% of the time while other critical lures (e.g. "butterfly") were never falsely recalled. Deese was able to predict the intrusion rates for the critical lures on the basis of the average associative strength from the studied associates to the critical lures and obtained a correlation of R=0.80. Because Deese could predict intrusion rates with word association norms, the WAS vector space derived from the association norms should also predict them. Critical items with high average similarity (or low average distance) to the list words in the semantic space should be more likely to appear as intrusions in free recall. The average similarity (average distance) was computed between each critical lure vector and list word vectors, and the correlations were computed between these similarities and observed intrusion rates.

The middle panel in Figure 2 shows the results. The pattern of results is quite different than the pattern of results for either recognition or cued recall. The best correlation of 0.82 was obtained with $S^{(1)}$, the sum of backward and forward associative strength. This result is very similar to the correlation of 0.80 Deese obtained with his word association norms. Interestingly, the plot shows that any manipulation that includes the indirect associations leads to worse performance than using the direct associations only. The WAS space based on $S^{(2)}$ now does worse than the WAS space based on $S^{(1)}$, and either space correlates more poorly than when using the associative strengths $S^{(1)}$ and $S^{(2)}$ themselves.

These findings imply that direct associative strengths are the best predictors of intrusion rates in free recall. One explanation for this finding is related to implicit associative responses (IAR's). Underwood (1965) has argued that during study, the words associated with the study words are thought of and might be stored in memory as an implicit associative response. In Deese's study, it is likely that IAR's were generated because the critical lures were all strongly associated to the list words. Therefore, during recall, the words that were actually presented and words that were thought of during study might be confused leading in some cases to dramatic intrusion rates. Because free associations measure what responses are first thought of given specific cues, the direct associative strengths can be argued to be good predictors of the strength of implicit associative responses and subsequent intrusion rates.

Discussion

By a statistical analysis of a large database of free association norms, the Word Association Space (WAS) was developed. In this space, words that have similar associative structures are placed in similar regions of the space. In the first version of WAS, singular value decomposition was applied on the direct associations between words to place these words in a high dimensional semantic space. In the second version of WAS, the same technique was applied on the direct and indirect associations between words. In the third version of WAS, metric multidimensional scaling was applied on measures for the associative strength related to the shortest associative path between words (similar to the approach in Cooke et al., 1986 and Klauer & Carroll, 1995).

Because the free association norms have been an integral part in predicting episodic memory phenomena (e.g. Deese, 1965; Nelson, Schreiber, & McEvoy, 1992), it was assumed that a semantic space based on free association norms would be an especially useful construct to model memory phenomena. We compared WAS with LSA in predicting the results of several memory tasks: similarity ratings in recognition memory, percentage correct in extralist cued recall and intrusion rates in free recall. In all these memory tasks, WAS was a better predictor for performance than LSA. This suggests to us that WAS forms a useful representational basis for memory models that are designed to store and retrieve words as vectors of feature values. Many memory models assume that the semantic aspects of words can be represented by collections of features abstractly represented by vectors (e.g. Hintzman, 1988; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997, 1998). However, in most memory modeling, the vectors representing words are arbitrarily chosen and are not based on or derived by some analysis of the meaning of actual words in our language. We expect that memory models based on these semantic vectors from WAS will be useful for making predictions about the effects of varying semantic similarity in memory experiments for individual words.

We propose that WAS is an approach that augments other existing methods available for placing words in a psychological space. It differs from the LSA approach in several ways. Because LSA operates on samples of text, it is relatively to apply LSA to large numbers of words. In contrast, the number of words that can be scaled by WAS depends on the number of words that can be normed. It took Nelson et al. (1999) more than a decade to collect the norms, highlighting the enormous human overhead of the method. Even though a working vocabulary of 5000+ words in WAS is much smaller than the 70,000+ word long vocabularies of LSA, we believe it is large enough for the purpose of modeling performance in variety of memory experiments. An advantage of LSA is the potential to model the learning process that a language learner goes through. For example, by feeding the LSA model successively larger chunks of text, the effect that learning has on the similarity structures of words in LSA can be simulated. In WAS, it is in principle possible to model a language learning process by collecting free association norms for participants at different stages of the learning process. In practice however, such an approach would not easily be accomplished. In any event, we believe that both WAS and LSA provide semantic spaces that are both useful for theoretical and empirical research.

The differences between the applications to different tasks certainly suggest that the usefulness of a particular semantic space will be task dependent. We speculate that the spaces differ in their semantic 'reach'. WAS is derived from the first associations provided, and thereby might emphasize local semantic domains, and such things as two-word units in memory. The fact that is does well for judgments of similarity to an episodic list might suggest that such judgments are based on a few episodic recalls cued by the test word, and that such recalls reflect 'nearby' associations. On the other hand the fact that WAS does better than the associations themselves suggests that recalls involve additional semantic components beyond those 'close by'. The fact that free recall intrusions favor associations over WAS suggests that this task favors very local semantics (perhaps through IARs). The fact that LSA lags behind the WAS based measures suggests that LSA captures semantics with a wider reach than WAS or the associations themselves. If so, it ought to be possible to find other tasks that would favor LSA, but such research has not yet been carried out.

References

Brainerd, C.J., & Reyna, V.F. (1998). When things that were never experienced are easier to "remember" than things that were. Psychological Science, 9, 484-489.

Bousfield, W.A. (1953). The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 49, 229-240.

Burgess, C., & Lund, K. (2000). 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.

Cooke, N.M., Durso, F.T., & Schvaneveldt, R.W. (1986). Recall and measures of memory organization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 538-549.

Deese, J. (1959a). Influence of inter-item associative strength upon immediate free recall. Psychological Reports, 5, 305-312.

Deese, J. (1959b). On the prediction of occurrences of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58, 17-22.

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

Derweester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., & Harshman, R. (1990). Indexing by latent semantic analysis. Journal of the American Society for Information Science, 41, 391-407.

Gillund, G., & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-67.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. Psychological Review, 95, 528-551.

Jenkins, J.J., Mink, W.D., & Russell, W.A. (1958). Associative clustering as a function of verbal association strength. Psychological Reports, 4, 127-136.

Klauer, K.C., & Carroll, J.D. (1995). Network models for scaling proximity data. In, R.D. Luce, M. D'Zmura, D. Hoffman, G.J. Iverson, & A.K. Romney (eds.), Geometric representations of perceptual phenomena. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Kucera, H., & Francis, W.N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

Landauer, T.K., Foltz, P., & Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes, 25, 259-284.

Marshall, G.R., & Cofer, C.N. (1963). Associative indices as measures of word relatedness: a summary and comparison of ten methods. Journal of Verbal Learning and Verbal Behavior, 1, 408-421.

McClelland, J.L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. Psychological Review, 105, 724-760.

Nelson, D. L. (2000). The University of South Florida Database on Cued Recall. http:// luna.cas.usf.edu/~nelson/

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida Word Association, Rhyme and Fragment Norms.http://luna.cas.usf.edu/~nelson/

Nelson, D.L., McKinney, V.M., Gee, N.R., & Janczura, G.A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. Psychological Review, 105, 299-324.

Nelson, D.L., Schreiber, T.A., & McEvoy, C.L. (1992). Processing implicit and explicit representations. Psychological Review, 99, 322-348.

Nelson, D. L., & Zhang, N. (2000). The ties that bind what is known to the recall of what is new. Psychonomic Bulletin & Review, 7, 604-617.

Nelson, D. L., Zhang, N., & McKinney, V. M. (2001). The ties that bind what is known to the recognition of what is new. Journal of Experimental Psychology: Learning, Memory and Cognition, 27, 1147-1159.

Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). The measurement of meaning. Urbana:

University of Illinois Press.

Shiffrin, R.M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. Psychonomic Bulletin & Review, 4, 145-166.

Shiffrin, R. M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.). Rational models of cognition. (pp. 73-95), Oxford, England: Oxford University Press.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. Journal of Experimental Psychology, 70, 122-129.