UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Is There Preferential Attachment in the Growth of Early Semantic Noun Networks?

Permalink

https://escholarship.org/uc/item/49q2x9kf

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN

1069-7977

Authors

Hills, Thomas T.
Maouene, Mounir
Maouene, Josita
et al.

Publication Date

2008

Peer reviewed

Is There Preferential Attachment in the Growth
of Early Semantic Noun Networks?

Thomas T. Hills (thomas.hills@unibas.ch)
[1]Department of Psychological and Brain Sciences, 1101 E. Tenth Street
Bloomington, IN, 47405 USA

Mounir Maouene (mmaoune@ensat.ac.ma) [1]
UFR : Artificial Intelligence and Bioinformatics, ENSAT, Tangier, Morocco

Josita Maouene (jcmaoune@indiana.edu) [1]

Adam Sheya (aasheya@indiana.edu) [1]

Linda B. Smith (smith4@indiana.edu) [1]

Abstract

Recent research on the statistical structure of semantic networks has suggested that the structure of the adult lexicon reveals evidence of a learning mechanism involving preferential attachment. With preferential attachment, words entering the lexicon make connections to existing words in proportion to the connectedness of the existing words. We asked if preferential attachment—or other growth processes—might be observed in children's early semantic networks and if so, what assumptions would be required for word relations. To do this, we used words from the MacArthur-Bates Child Development Inventory over a 15-month period (16 to 30 months)—and the word acquisition patterns of 20 individual children—to build longitudinal semantic networks using two separate assumptions for semantic relatedness. These involved edges based on feature similarity using the McRae et al. (2003) feature norms and edges based on association using the University of Southern Florida Free Association Norms. The feature-based network showed no evidence of preferential attachment, or the opposite, preferential avoidance. However, the association-based network did show preferential attachment, even across this very small time span. Currently, we are investigating in more detail the growth of this network, and other possible growth mechanisms besides preferential attachment, in an effort to better understand the mechanisms that drive early semantic growth.

Keywords: Early semantic networks, growing network models, word learning, preferential attachment.

Introduction

Word learning begins sometime prior to a child's first birthday. It is very slow at first with each word added seeming effortful and with false starts (Nelson, 1973; McNamara, 1982). But 6 months later, early word learning shows an accelerating function (Bloom, 2000). At first, most of these new additions are common nouns, the focus of this paper. Developmentalists have studied these early noun vocabularies interested both in the kind of categories that start the system off and how early noun learning might create the needed knowledge to facilitate the ensuing accelerated learning rate (Colunga & Smith, 2003). Although it seems reasonable that later noun learning builds on earlier noun learning, it need not be. Tomasello (2000) has argued—with respect to syntactic learning—that children's earliest words are part of an imitative, item-by-item process, in which word learning is not dependent on prior word knowledge—that is, the addition of new words is independent of the existing lexicon. Similar rules may apply to early semantic acquisition.

The primary goals of this work involve asking two questions. First, what are meaningful ways to represent the development of children's early semantic networks, prior to 30 months of age? To answer this question, we take two approaches. One—feature relations—is based on a host of prior work in the developmental literature showing that children recognize object features and can use feature information to identify and place objects into higher-order categories (e.g., Landau, Smith & Jones, 1988; Bomba & Siqueland, 1983; Gelman & Bloom, 2000). In many studies like these, e.g., Booth & Waxman (2002), children generalize object names across categories with respect to perceptual or functional features. The children's feature network we investigate here has also been shown to produce meaningful object categorizations consistent with adult knowledge (Hills et al., 2008). In principle, the generalization process associated with object categorization creates opportunities for preferential word learning by allowing children to differentiate known objects through processes of word production (or expectation) and subsequent error correction.

The second—relations based on free association—has seen dramatic successes in predicting adult performance in semantic decision and memory tasks (e.g., Griffiths et al., 2007; McEvoy, Nelson, & Komatsu, 1999) and recent success in predicting age of acquisition (AoA) at grosser temporal scales than we address here (e.g., binning data between 12 and 132 months before and after 52 months), and using only data from the full adult network (Steyvers & Tenenbaum, 2005). In the present work we take a closer

look at the growth process of this network by following its development in children.

This leads to our second question: how do these early semantic networks grow? Recent work using network analyses to investigate the structure of large-scale adult semantic networks provides one new approach to understanding how (whether and when) children's past word knowledge facilitates the acquisition of new and semantically related words. This work suggests that semantic networks may grow according to item-dependencies, in which new semantic relationships are formed based on the structure of the existing semantic network (Steyvers & Tenenbaum, 2005).

The network approach follows an early proposal by Quillian (1967) that memory access comes about via semantic relatedness. Collins and Loftus (1975), Anderson (1976), and Posner & Snyder (1975), developed these ideas further into models of spreading activation, in which words were represented by nodes and relationships between words were represented by links (or edges). In models of spreading activation, activity at one node in the network is assumed to bias the activity of other nodes via the connections between them.

Steyvers & Tenenbaum (2005) investigated the statistical structure of large-scale semantic networks by combining words with inter-word relations based on evidence for semantic relatedness. Specifically, they constructed three large-scale semantic networks—generating edges between words using the University of Southern Florida Free Association Norms, WORDNET, and Roget's Thesaurus—and found robust statistical patterns of connectivity, which included small world structure and power-law structure in their degree distributions.

Power-law structure can be a consequence of preferential attachment, in which words entering the lexicon make connections to existing words in proportion to the number of relations (degree) of the existing words (Barabasi & Albert, 1999). This creates a rich get richer phenomenon found in a variety of network structures. Steyvers & Tenenbaum (2005) built a model of preferential attachment in semantic networks that generated power-law structure, based on the assumption that words are added to the network to facilitate the differentiation of complex (high-degree) words that are already known—similar to the feature network differentiation process we describe above.

If words are added to early semantic networks based on rules associated with preferential attachment, then words that are learned earliest in these networks should show proportionally more connections during later stages of network development. Steyvers & Tenenbaum (2005) confirmed this for associative relations using adult estimates of age-of-acquisition (Gilhooly and Logie, 1980) and age of acquisition norms based on when children can name an object in the laboratory (Morrison, Chappell & Ellis, 1997).

In the present work we ask if these same patterns of connectivity are observed in association-based and feature-based networks using words known prior to 30 months. Our analysis also asks if preferential attachment patterns are observed when the networks are observed in a month-by-month fashion—or if other patterns are observed.

The Shared Feature Network

The nouns. The 130 noun categories were selected from the Bates-MacArthur Communicative Developmental Inventory or Bates-MCDI (Fenson et al., 1994), Toddler version. This inventory is a checklist of 682 words of which 285 are basic-level nouns (excluding body parts). These words are the first words typically acquired by children learning English and normatively include in the productive vocabularies of 50% of children at 30 months of age. We used the month-by-month norms of the percentage of children with these words in their productive vocabulary to construct a network for each month from 16 to 30 months. We let the age of acquisition for a word be equal to the first month at which the word is produced by more than 50% of the children in the MCDI. The checklist contains 285 basic level nouns (excluding body parts). The nouns studied here are a subset consisting of the nouns on the Bates-MCDI that were also included in the McRae et al. feature generation studies. These 130 nouns over-represent animals (33 nouns, 25% of the subset versus 43 nouns, 15% of the Bates-MCDI nouns) and under-represent food (17 nouns, 13% of the subset versus 68 nouns, 23% of the Bates-MCDI nouns). Nonetheless, the sample includes a broad array of nouns across relevant superordinate categories.

The features. The features associated with each of the 130 nouns were taken from the feature norms reported by McRae, Cree, Seidenberg, & McNorgan (2003). That study collected feature norms for 541 concepts from a total of 725 adult participants with 30 participants providing features for each concept. The participants in that study were given each noun and 14 blank spaces to fill with features and were prompted to provide physical properties (how it looks, smells, sounds etc), functional properties or uses, internal properties, and other pertinent facts. The generated features were also classified by the Cree and McRae's (2003) theoretical brain region taxonomy into the following categories: functional, visuo-motor, visual form, visual color, sound, taste, smell, tactile, encyclopedic, and taxonomic. For the present study, we only considered the functional and perceptual features excluding encyclopedic and taxonomic features as unlikely to be available to children's direct experience.

The network. To construct the developing network used for the subsequent analyses, we let nouns represent nodes. Then, for every noun-pair, the value of the edge between them represents the number of their shared features, using the McRae feature norms. For example, BALL and APPLE will share an edge of weight one because they both share the feature IS_ROUND. An edge between two nodes will be taken as the addition of one degree in connectivity, regardless of the weight of the edge. In Figure 1 we present the noun network at 30 months derived from creating edges between nouns based on shared features where, for visual

156

clarity only edges representing two or more features are drawn.[1]

Figure 1: Noun network for children at 30 months of age based on shared features. The network has 130 nodes with 824 edges (not all are shown), with a density of approximately 0.1. The average degree in the network is 12.67, with a clustering coefficient of 0.53. More statistics are provided in Hills et al., 2008.

Results

The Feature-based Network

As an initial test of the hypothesis that words learned during the earliest stages of development have more connections during later stages—consistent with a model for growth with preferential attachment—we plotted the mean degree for each word in the 30-month network against the age of acquisition for the word (Figure 2). The figure provides no evidence that words learned earlier are more likely to have higher connectivity during later stages of development (up to 30 months).

Another signature of growth with preferential attachment is that the degree connectivity of the network has a power-law distribution. This is revealed by a linear pattern when the degree distributions of the nodes are plotted on a log-log plot. Figure 3 shows the log-log plot for the cumulative degree distribution of the 30-month network. The y-axis presents the probability density for finding a node of greater than or equal to degree k. The resulting distribution is more consistent with the exponential distribution presented for non-power-law networks in Steyvers & Tenenbaum (2005). In other words, unlike the USF Free Association, WORDNET, and Roget networks, the early feature-based

---

[1] All results shown here for the feature network were unchanged by building the network at different thresholds for inclusion of an edge (e.g., greater than 1 shared feature, greater than 2 shared features, or up to greater than 3 shared features). Beyond which the network is too sparse for meaningful analysis.

network does not appear to have the structural properties consistent with preferential attachment when observed at 30 months of age.

Figure 2: Mean degree by age of acquisition for nouns learned during the first 30 months of development based on the shared feature network in Figure 1. Error bars represent standard error of the mean.

Figure 3: The log-log plot of the cumulative degree distribution for the 30-month- feature network.

Because we have month-by-month growth data for our semantic network, we are able to evaluate the growth patterns in more detail. Specifically, we asked if the feature networks might grow according to a mechanism besides preferential attachment. In the world of features, objects may become salient for the very reason that they do not generalize to known objects via shared features—in which case, once one knows the word DOG, it may be easier to learn the word TELEPHONE than CAT. We looked at the probability of a word acquiring a new feature-based relation (a new edge) for each month that the word was in the network. To do this, we took the degree of each node at the start of the month and evaluated how many new edges the node received by the addition of words in the following month. To make sure we are obtaining a cross-section of words with sufficient opportunity to receive new

connections, we only consider words with more than 10 opportunities (i.e., months) to acquire a new edge. Results are presented in Figure 4.
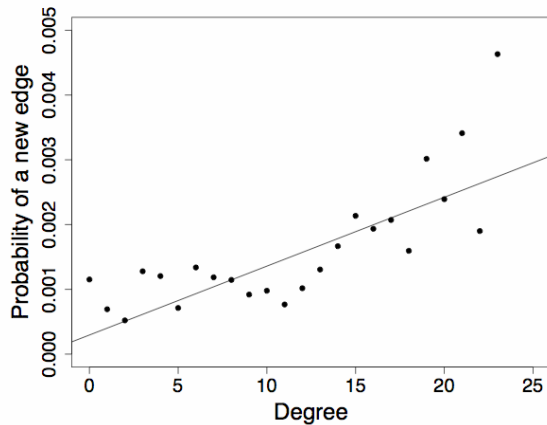

Figure 4: Probability of acquiring a new edge as a function of the nodes degree (k) in the previous month.

Figure 4 reveals a significant positive relationship between the degree of a node and the likelihood the node will receive a new semantic edge in the following month (R^2 = 0.59, p < 0.001). Words with more semantic relationships are more likely to acquire new relationships in subsequent months. This suggests the possibility that preferential attachment *is* playing a role in our network's growth, but—given the prior evidence presented in Figures 2 and 3—it also suggests that degree is not a causative factor in the growth of the network. In other words, it may be that the statistical structure of the edge addition probabilities is no different from what would be observed if the networks were generated by randomly acquiring words. To test this, we generated 100 simulated network growth sequences by randomizing the age of acquisition for words in each of these networks. Figure 5 shows that the mean regression line and the 95% confidence intervals for the simulated data (dotted lines) are a very close fit to the regression line for the observed data (solid line). The degree of words in the network does not appear to influence the acquisition of new words.

One caveat associated with using the MCDI data is that the word acquisition data is normative, and therefore not representative of individual children, but simply of an *average* child. It may be every individual child shows evidence of better than random preferential attachment in the feature-based network, but that once children are averaged together to construct the age of acquisition norms, the evidence for this pattern disappears. We tested this by collecting (through a longitudinal study) age of acquisition data for the 130 nouns for 20 individual children (from 18 to 30 months of age) and then for each child independently calculating the relationship between degree of a word and the probability of acquiring a new edge in the subsequent month. The thick dotted line in Figure 6 shows the results of these 20 individual regression lines averaged together. Here again, the evidence for preferential attachment is no

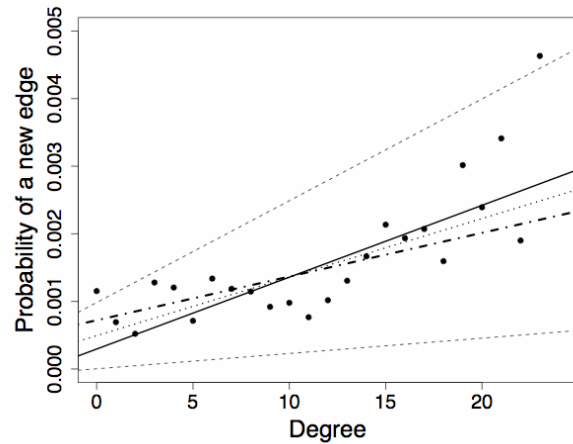better than that predicted by a random addition of words to the network.

Figure 5: Probability of acquiring a new edge as a function of the nodes degree (k) in the previous month. Thin dotted lines show the mean and 95% confidence intervals for 100 simulated network growth sequences, in which words are learned in a random order. The thick dotted line shows the average regression line for a sample of 20 individual children. The solid line is the same as in Figure 4.

The Free Association Network

If young children are truly item-by-item learners with words learned independently of one another, then the same results should be obtained with an association network—a network for which degree does correlate with AoA at a much grosser level of analysis (e.g., when words between 12 and 132 months are binned before and after 52 months; Steyvers & Tenenbaum, 2005). This may be perfectly reasonable if the environments children learn their earliest words from—e.g., child-directed speech and the objects presented to children during play—do not invoke the same structural patterns as adult reported associations (or features). On the other hand, associations may capture a more constrained measure of relatedness than features—because the feature norms ask for multiple responses per subject (as compared with one response for associates) and they also elicit category general features, such as IS_DIFFERENT_COLORS and IS_DIFFERENT_SIZES. To examine these issues, we extracted from the USF Free Association data (Nelson, McEvoy & Schreiber, 1999) the network based on associative relations using the same 130 nouns used for the feature-based network. Figures 6 shows what the network looks like at 30 months. We then evaluated the degree at 30 months with the age of acquisition recorded by the MCDI.

Contrary to what we found for the shared feature networks, the network based on adult free associations does show a pattern consistent with preferential attachment—with earlier acquired words showing higher degree at 30 months of age (Figure 7). The log-log plot presented in Figure 8 is also consistent with this pattern, with a

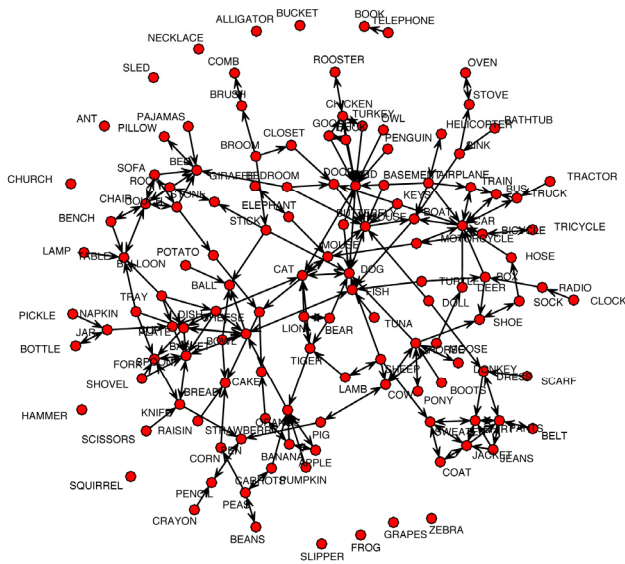substantially more linear pattern than that presented in Figure 3.

Figure 6: Noun network for children at 30 months of age based on direct association in the USF Free Association Norms. Data from the Free Association Task are shown here with arrows pointing from stimulus (cue) to the response (target). All targets were produced by two or more subjects (as described in Nelson et al., 1999).
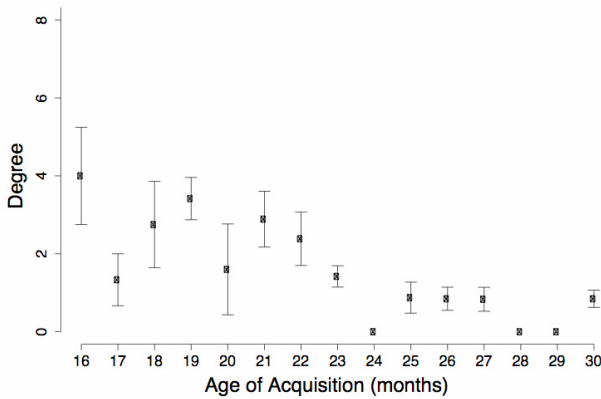

Figure 7: Mean degree by age of acquisition for nouns learned during the first 30 months of development based on the free association network in Figure 6. Error bars are standard error of the mean for words acquired at a given age.

Discussion

In the present work we investigated the possibility that early noun learning reveals a more-gets-more pattern of development, with words of higher semantic connectedness being more likely to acquire new semantic relations during lexical development. Our results suggest that, unlike the three adult networks presented in Steyvers & Tenenbaum (2005), when the structure of children's semantic information is based on feature-resemblance, there is no evidence for preferential attachment (or its opposite

preferential avoidance). However, if we ask a similar question of the early semantic network formed by creating edges based on adult free association, then the network does show evidence for preferential attachment.
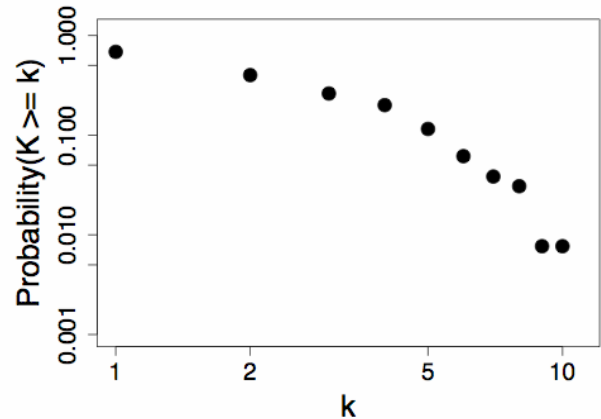

Figure 8: The log-log plot of the cumulative distribution for the free association network.

Steyvers and Tenenbaum (2005) point out that some semantic resources do not show the statistical structure associated with preferential attachment without special assumptions. For example, Latent Semantic Analyses (LSA) (Landauer & Dumais, 1997), when used in a standard way (using cosine similarity) to generate relations between words, generates a degree distribution qualitatively similar to that shown in Figure 3 (shown in Figure 8 of Steyvers & Tennenbaum, 2005). Evidence for preferential attachment is strongly dependent on one's choice of semantic relatedness. The lack of any evidence for a statistical pattern predicting AoA in the feature network suggests that adult-generated features may not be important in the acquisition of new words. However it is poses an interesting question about the relation between feature-based and association-based semantics. Others have pointed out that adult associations are often paradigmatic, and feature associations are very likely to conserve that relation. However, free associations may benefit from a more constrained retrieval process, because only one associate is given for each cue. In the feature norms we use here, multiple features are listed, and this may wash out salient features that drive associations (in those places were associations are feature driven). We are currently looking at production frequency in the feature norms to parcel out this possible relation.

Associations, on the other hand, are predictive and deserve further investigation in developmental studies.

Finally, finding power-law structure in adult networks is not the same thing as finding growth with preferential attachment. While it is certainly true that one can generate power-law structure with models based on preferential attachment (e.g., Barabasi et al., 1999), random sampling from statistical structures that already have a power-law structure will arrive at the same place. For example, the cognitive processes associated with the evolution of language may generate power-law semantic structure, and

children may simply sample words from this adult structure—with words of higher degree being sampled at earlier ages. The evidence provided here and elsewhere does not provide a clear causative direction. However, we can ask more specific developmental questions about the growth of these networks—especially with respect to hypotheses that may produce similar semantic structure as preferential attachment in the adult state. One model we are currently investigating is based on preferential *acquisition*, in which words are added to the network based on their degree in the adult state. Analysis of specific preferential attachment models suggest that preferential acquisition models are at least as viable a developmental pathway as preferential attachment and, and in some cases, explain aspects of the network growth that preferential attachment does not (Hills et al., in prep).

Acknowledgments

References

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286,* 509-512.

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

Bomba, P.C., & Siqueland, E.R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology, 35*, 294-328

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Colunga, E. & Smith, L. B. (2003) The emergence of abstract ideas: evidence from networks and babies. *Philosophical Transactions by the Royal Society,* 358, 1205-1214.

Cree, G. S. and McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of Chipmunk, Cherry, Chisel, Cheese, and Cello (and many other such concrete nouns). *Journal of Experimental Psychology: General,* 132, 163-201.

Gelman, S.A., & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding how to name it. *Cognition, 76,* 91-103.

Gilhooly, K. J., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behavior Research Methods and Instrumentation, 12,* 395-427.

Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind. *Psychological Science, 18,* 1069-1076.

Hills, T., Maouene, M., Maouene, J., Sheya, A. & Smith L. B. (2008). Categorical Structure in Early Semantic Networks of Nouns. *Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC, USA, July 23-26.*

Hills, T., Maouene, M., Maouene, J., Sheya, A. & Smith L. B. (In prep) Early semantic network growth: preferential attachment versus preferential acquisition.

Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development 3(3)*, 299-321.

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.

McNamara, J. (1982). *Names for things: A study of human learning.* Cambridge MA: MIT Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavioral Research Methods, Instruments, and Computers, 37*, 547–559.

Morrison, C. M., Chappell, T. D., & Ellis. A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A*, 528-559.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*(1-2), No 149.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of Southern Florida word association norms.* Retrieved from http://w3.usf.edu/FreeAssociation

Posner, M. I., & Synder, C. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* . London: Academic Press.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Sciences, 12*, 410–430.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science, 29*, 41-78.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences 4(4)*, 156-163.

Waxman, S. & Markow, D. (1995). Words as invitations to form categories: Evidence from 12-month-old infant *Cognitive Psychology, 29,* 257-302.

Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Child Development, 54,* 858–867.