UC Merced

Title
Mutual Exclusivity and Vocabulary Development

Permalink
https://escholarship.org/uc/item/8sr8456n

Journal

ISSN

Authors
Yurovsky, Daniel
Bion, Ricardo
Smith, Linda
et al.

Publication Date
2012

Peer reviewed

Mutual Exclusivity and Vocabulary Structure

Daniel Yurovsky
dyurovsk@indiana.edu
Department of Psychological and Brain Sciences
Indiana University
Linda B. Smith
smith4@indiana.edu
Department of Psychological and Brain Sciences
Indiana University

Ricardo A. H. Bion
ricardoh@stanford.edu
Department of Psychology
Stanford University
Anne Fernald
afernald@stanford.edu
Department of Psychology
Stanford University

Abstract

The words that children learn can be characterized as a semantic network, with links connecting related words. Recent analyses have shown these networks to have small-world structure, with a few highly-connected hub words facilitating short paths between otherwise distant words. This structure contributes to network robustness, and differences in structure can predict differences in language learning outcomes. While previous studies have shown that semantic network structure reflects linguistic input structure, we provide the first evidence that it is related also to children's own language learning biases. Two-year old children who show a mutual-exclusivity bias have significantly more hub-like networks than children who do not, even when they know the same number of words. This finding contributes to our understanding of both semantic networks and the origins of mutual exclusivity.

Keywords: word learning; mutual exclusivity; semantic networks; language acquisition

Introduction

Although the earliest analyses of human memory and learning concerned the learning of lists of unrelated words (Ebbinghaus, 1885/1962), researchers quickly discovered that the words people learn in more natural contexts are intricately connected. Vocabularies were conceptualized as richly structured networks, with links connecting semantically related words (Collins & Loftus, 1975). These connections play an important role in both learning and memory, and can be observed empirically in semantic priming experiments. Because activation spreads from words to their semantic neighbors, presenting a word, even subliminally, leads to faster processing of related words (Anderson, 1983). Even two-year old infants show semantic priming, suggesting that vocabularies have network structure early in language learning (Arias-Trejo & Plunkett, 2009).

Recently, the application of graph-theoretic methods to the study of these networks has begun to provide insight into their structural properties. For instance, Hills, Maouene, Maouene, Sheya, & Smith (2009a) analyzed the semantic network structure of 130 nouns typically learned before 30 months. Compared to randomly-connected control networks, these semantic networks showed significant small-world structure, in which most words are sparsely connected, but a few are highly-connected hubs. This kind

of structure results in networks robust to malfunction (e.g. forgetting a word; Albert, Jeong, & Barabási, 2000), and can help to explain some of the remarkable efficiency of human semantic memory (Raaijmakers & Shiffrin, 1981). Further, semantic networks lacking this structure are associated with slower language-learning, characterizing the vocabulary structure of late talkers (Beckage, Smith, & Hills, 2011). But why do children learn these words? Why do semantic networks have this structure?

Undoubtedly, one answer to this question is that structure comes from the environment. Because children learn words from the language they hear, language input is a strong predictor of the words that children will learn. For instance, the frequency with which a child hears a word in isolation can predict how likely a child is to learn that word (Brent & Siskind, 2001). Similarly, the semantic networks constructed from corpora of both adult-directed and child-directed language have many of the same structural properties as networks constructed from the words 30-month-old children are likely to know (Hills, et al., 2009a; Steyvers & Tenenbaum, 2005).

But perhaps a more complete explanation of the origin of semantic network structure is that it emerges from an interaction between structure in the linguistic environment and the child's own learning system. Because children are not unbiased samplers of linguistic input, their attentional and learning biases mediate the link between language input and language learned (Hudson Kam & Newport, 2005; Smith, 2000). For instance, children who learn to attend to shape are likely to learn shape-based categories, and those who learn to attend to other properties (e.g. material) learn other kinds of words (Colunga & Sims, 2011; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Can word-learning biases predict and explain semantic network structure? In this paper, we consider the case of disambiguation through mutual exclusivity.

In the disambiguation task, a child is presented with a novel object among one or more familiar object competitors. The child then hears a novel label (e.g. 'can you find the dax?') and is asked to select an object. Both toddlers and adults reliably select the novel object as the target of the novel label (Markman & Wachtel, 1988; Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992), and studies with infants suggest that this disambiguation may arise as early as 18 to 22.5 months (Halberda, 2003; Mather & Plunkett, 2009). Preferential

mapping of novel labels to novel objects over known objects, which we will refer to as mutual exclusivity (ME), could arise for a number of reasons, and its mechanism of action is the topic of significant debate (e.g. Diesendruck & Markson, 2001; Golinkoff et al., 1992; Markman & Wachtel, 1988). We explore this question in the general discussion, but will sidestep it here and instead consider the potential consequences of mutual exclusivity for semantic network structure.

Mutual exclusivity is a mechanism by which children can leverage prior knowledge to learn new words *in the context of known objects*. Consequently, children who show mutual exclusivity should have vocabularies that echo this kind of contextual structure. For these children, learning *fork* should ease the acquisition of *spoon*, *bowl*, and *plate*. In contrast, learning *fork* should have little effect on the acquisition of *dog* and *coat*. Thus, we propose that mutual exclusivity can help explain small-world structure of semantic networks, and those children who show mutual exclusivity will have more hub-like networks than those who do not. We begin by reporting empirical data from a disambiguation task with 24-month-old children, continue by describing a semantic network analysis of these children's vocabularies, and conclude with a discussion of how these results inform our understand of the relationship between mutual exclusivity and vocabulary development, as well as the origins of mutual exclusivity itself.

Experiment

Method

Participants. Forty two-year-olds ($M = 24.75$ months; range = 24-26; 20 female) participated. All were typically developing children from households in which parents reported English to be the dominant language. A subset of 34 infants ($M = 24.9$ months; range = 22.4-27.5; 16 female) participated in the followup analysis (explained below).

Stimuli. Nine familiar objects (e.g. boat, glasses) were used in the warm-up trials. Twenty-five familiar (e.g. brush, cup) and 8 novel objects (e.g. massager, platypus) were used in the referent-selection task.

Procedure. Parents first completed the MCDI (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994) and an SES measure (Hollingshead, 1975). After this, each child participated in three warm-up trials. On warm-up trials, the experimenter set a tray containing three familiar objects on the table, initially covered by an occluder. The experimenter asked for the target object (e.g. "which one is the dog?") three times: once while the items were occluded, again after lifting the occluder, and again three seconds later while pushing the tray towards the infant. The first reach, point, or grab, was scored as a response. On these trials, infants were praised for correct responses and corrected when necessary.

Subsequently, each child participated in sixteen referent-selection trials. On each trial, the experimenter presented a tray containing two familiar objects and one novel object.

The procedure was identical except that children received neutral feedback on all trials. On half the trials, the experimenter asked for a familiar object, while on the other half she asked for a novel object (e.g. *modi*, *taju*).

Results and Discussion

Each participant made a total of 16 choices, picking 8 targets on familiar trials, and 8 targets on novel trials. Any trial on which the child did not know the label for the familiar target, or the label for one of the familiar distractors, was excluded from analysis. The proportion of targets correctly chosen on these remaining trials was then analyzed to determine the child's success in the task. Overall, children performed quite well, selecting the correct target on both familiar ($M_f = .83$, $t(39) = 15.31$, $p < .001$) and novel trials ($M_n = .545$, $t(39) = 5.87$, $p < .001$) at greater than chance levels. Thus, as a group, 24-month-old children used mutual exclusivity for disambiguation. Familiar trial performance, however, was significantly higher than novel trial performance ($t(39) = 6.88$, $p < .001$).

Because the central question in this study is about the relationship between learning mechanisms and vocabulary development, we measured both vocabulary size (MCDI - Fenson, et al., 1994) and mother's education (Hollingshead, 1975), a potential correlate of rich language input. Mother's education was reliably correlated with performance on familiar trials ($r = .33$, $p < .05$), but not novel trials ($r = .01$, *n.s.*), and vocabulary size was not significantly correlated with performance on either kind of trial ($r_f = .19$, *n.s.*; $r_n = .15$, *n.s.*). In the semantic network analysis to follow, we show that *vocabulary structure* is reliably related to novel trial performance. Because neither mother's education nor *vocabulary size* predict ME in this data set, the relationship between ME and structure is likely to be quite robust.

But perhaps this analysis is unfair. While most of the children had high levels of success on familiar trials, a few children did not perform as well. Since these children knew the words for all three objects on these familiar trials, their low levels of performance indicate that they may not have understood the task. Thus, for the same reason that response time analysis typically uses only correct response trials, excluding these children from individual-level analyses may give clearer correlations. In order to determine whether a child's performance was significantly better than expected by chance, we modeled chance behavior on each trial as random selection of one of the three objects.

The probability of success expected by chance is given by a binomial distribution with probability ⅓. Consequently, a child should be counted as performing differently from chance if he or she made enough correct selections to be outside the 95% confidence interval for a binomial distribution. A child who made 8 choices, for instance, needed to make at least 5 correct choices to be counted as performing better than expected by chance. Each child's number of correct selections on familiar trials was thus submitted to a binomial test. Six of the 40 children were found to have performance levels on the familiar trials

indistinguishable from chance, and were thus excluded from further analysis. This left a subset of 34 children who could confidently be assumed to have understood the task. Figure 1 shows novel and familiar trial performance for children both from the full set, and from this reduced subset.

We also performed a similar analysis on novel trials, dividing children into two categories: those who reliably showed evidence of using mutual exclusivity (ME), and those who did not. Seventeen children were classified as ME users, and seventeen were classified as Nonusers. We are not arguing that ME is a binary phenomenon, but rather perform this binary split for technical reasons. Binarization loses some information separating children within the ME users category, but it also cleans up noise that may not meaningfully separate nonusers. Quantitative differences at or below chance levels are more likely to be generated by noise than they are to be generated by meaningful process differences, and thus are likely to dilute linear correlations. In subsequent analyses, because mutual exclusivity is analyzed as a binary phenomenon, we use Spearman's ρ, a non-parameteric measure of correlation. In all cases, correlations were stronger for this binary measure.

In this subset, mother's education was still correlated with performance on familiar trials ($r = .36$, $p < .05$), as was vocabulary size ($r = .39$, $p < .05$). Neither mother's education nor vocabulary size predicted performance on novel trials ($\rho = -.11$, $n.s.$; $\rho = .08$, $n.s.$). In the analyses that follow, we compare the semantic network connectivity of ME users and nonusers. Because use of mutual exclusivity was uncorrelated with vocabulary size, differences in network connectivity are unlikely to be a simple reflection of network size. Further, because mother's education predicted performance on familiar, but not novel, trials, a relationship between ME and vocabulary structure arising from language input must come from more specific properties not indexed by mother's education in this sample.
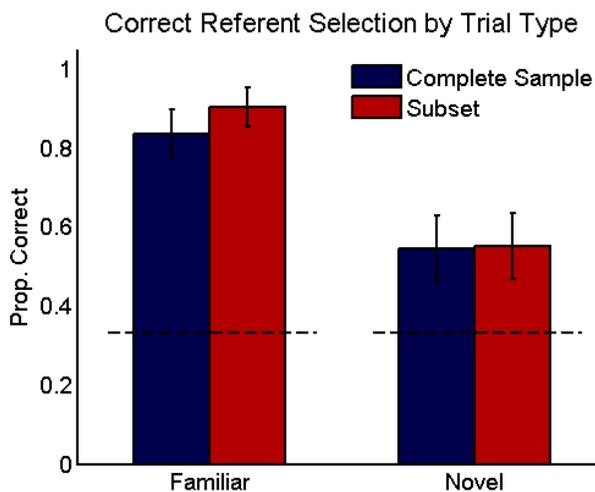
Figure 1: Proportion of correct choices by participants in both the familiar and novel conditions. Dark blue bars show the complete sample, light red bars the subset. Error bars indicate +/-1 standard error.

Semantic Network Analysis

To understand how use of mutual exclusivity contributes to the structure of children's vocabularies, we formalize these vocabularies as semantic networks. In semantic networks, vertices represent the words that children know, and edges represent semantic relationships among these words. In any such analysis, the first step is to formalize 'semantic relatedness' – the relationship used to link two words.

Previous analyses have used a number of successful metrics of connectivity: co-occurrence in CHILDES (e.g. Beckage, Hills, & Smith, 2011), frequency of free-association by adults (e.g. Steyvers & Tenenbaum, 2005), and common perceptual and conceptual features (e.g. Hills, et al., 2009b). In our analysis, we adopt and extend the last approach, connecting two words if they share a number of common semantic features. Features were drawn from the set of McRae feature norms (McRae, Cree, Seidenberg, & McNorgan, 2005). McRae and colleagues asked 725 adults to freely list up to 14 features of 541 English nouns. The number of features shared by two words gives a measure of their semantic relatedness.

Although participants could generate any features they liked, McRae et al. (2005) subsequently divided the generated features into 4 categories: perceptual features accessible to the 5 senses (e.g. "has fur," "tastes sweet"), functional features (e.g. "used for writing," "is edible"), encyclopedic features (e.g. "is expensive"), and taxonomic features (e.g. "a crustacean"). Following Hills et al. (2009b), we analyze only features of the first and second kind, as these are the features likely to be available to two-year-old children. We create two different networks for each child: one in which connectivity is defined by *perceptual* feature overlap, and one in which connectivity is defined by *functional* feature overlap. This is because overlapping perceptual features indicate a very different kind of relatedness than overlapping conceptual features.

Hills et al. (2009b) analyzed the clusters produced by each of these kinds of networks to quantify these different kinds of relatedness. Defining connectivity by *perceptual* feature overlap produced networks that were dense, highly connected, put words into more than one category, and produced categories that were overly inclusive relative to human judgments (e.g. MCDI categories, Fenson, et al., 1994). In contrast, *functional* feature overlap produced networks that were sparser, had smaller, better defined categories, and were better at discriminating among near-category members. In general, words connected in the *functional* network are more likely to be encountered in a relational context, facilitating learning by mutual exclusivity (e.g. *cake-carrots*, *boots-coat*). In contrast, words connected in the *perceptual* networks are less likely to be encountered in such situations, and learning one is thus less likely to facilitate learning the other through mutual exclusivity (e.g. *sheep-sofa*, *pencil-stick*). Thus, we can test a specific prediction about how mutual exclusivity builds vocabulary structure: it facilitates the acquisition of *functionally* related words.
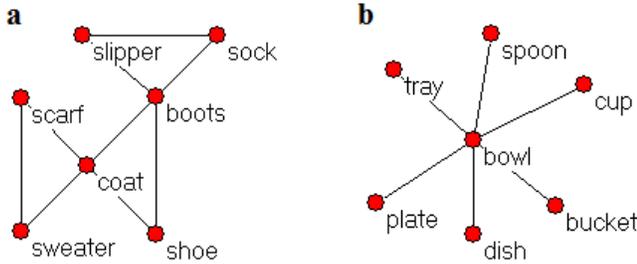
Figure 2: Two 7-vertex networks with different connectivity structures. The left network has a high clustering coefficient and a low degree centrality. The right network has a low clustering coefficient, but a high degree centrality.

In addition to using these two kinds of features to define connectivity, we measure their resulting structure in two different ways. These different connectivity measurements represent different ways in which mutual exclusivity could build structure. Consider the networks in Figure 2.

The first network (2a) has many local clusters, triangles in which any vertices with a common neighbor are likely to be neighbors themselves. One might predict mutual exclusivity to facilitate this kind of structure because using one word (e.g. *scarf*) to learn a semantic neighbor (e.g. *sweater*) should make a common neighbor even easier to learn (e.g. *coat*). This structure is measured by *clustering coefficient* (Equation 1), which has previously been used to distinguish the vocabulary structures of early and late talkers (Beckage, Hills, & Smith, 2011).

$$C = \frac{1}{|V|} \sum_{i=1}^{|V|} \frac{2|\{e_{jk}\}|}{d(v_i)(d(v_i) - 1)} \quad v_j, v_k \in N_i, e_{jk} \in E \quad (1)$$

In contrast, the second network (2b) does not have any local clusters, but rather has a single highly salient *hub*: a single vertex with many neighbors. This kind of structure might be even more likely to arise through mutual exclusivity, as learning the *hub* word (*bowl*) makes each of its neighbors easier to learn (*spoon*, *tray*, *cup*). This kind of structure is measured by *degree centrality* (Equation 2). This measure is new to semantic network analyses, but is a mainstay of social network science (Freeman, 1979), and measures the structural property intuitively most likely to be related to learning words through exclusion.

$$D = \frac{\sum_{i=1}^{|V|}[d(v^*) - d(v_i)]}{(|V| - 1)(|V| - 2)} \quad (2)$$

Thus, we test two hypotheses in the following analysis: mutual exclusivity should predict connectivity structure in *functional* but not the *perceptual* networks, and it should manifest more strongly in high *degree centrality* should than in *clustering coefficient*.

Method

To construct semantic networks for each child, we used all words which are both measured by the MCDI, and for which McRae and colleagues collected feature norms. This resulted in a list of 130 nouns, encompassing animals, food, clothing, vehicles, etc. For a full list, see Hills et al. (2009b). Each child's semantic network was constructed by adding one vertex for each word on that child's productive MCDI. Vertices were connected if they shared a minimum number (w) of semantic features. To be consistent with Hills et al. (2009b), we set this features threshold to all possible values 1-4. At $w = 3$, for instance, two words were connected only if they shared three or more semantic features. However, networks become increasingly sparse as w increases, and we thus urge caution in interpreting results at high thresholds.

Two networks were created for each child, one network in which only *perceptual* features defined connectivity, and one network in which only *functional* features were used to define connectivity (see above). Networks were defined by their set of vertices V and the set of edges E that connected them. A vertex's degree ($d(v)$) is defined as the number of other vertices to which it is connected by an edge. These connected vertices are called neighbors, and together define a node's neighborhood (N).

Once each network was constructed, two properties of its connectivity structure were measured. The first, clustering coefficient (C), measures the proportion of vertices with a common neighbor that are also neighbors of each other. (Equation 1). The second, degree centrality (D), measures the proportion of edges connected to a single dominant *hub* vertex (Equation 2). These measures of structure trade off, with high degree centrality necessitating a low clustering coefficient. Both measures always range between 0 and 1, and thus are independent of the size of a child's vocabulary. They are measures of structure independent of size.

Results and Discussion

As in the analysis above, children were divided into two groups: Mutual Exclusivity Users who performed better than chance on the novel trials of the disambiguation task, and Nonusers who did not. Again, we reiterate that this is not a theoretical commitment, but rather a tool for noise reduction. The structure of each child's individual semantic networks – both perceptual and functional – was used to predict that child's category of mutual exclusivity usage.

Before presenting the results of network analyses, we recapitulate that vocabulary *sizes* were quite comparable between these groups. The 17 ME Users produced an average of 408.3 words on the MCDI while the 17 Nonusers produced an average of 388.1 ($t(32) = .37$, *n.s.*). They also did not differ in the number of words they knew from the set of 130 used in the network analysis ($M_u = 92.6$, $M_n = 88.1$, $t(32) = .51$, *n.s.*). However, the particular words they knew, and the semantic relationships among them, proved to be importantly different.

Figure 3 shows correlations between measures of network structure and the mutual exclusivity category to which each individual child belonged. For perceptual networks, constructed by connecting words by shared perceptual features (e.g. "has fur," "tastes sweet"), neither clustering coefficient nor degree centrality were related to use of mutual exclusivity at any feature overlap threshold (Figure 3, left column). As predicted, perceptual networks, in which connections are not a good proxy for words likely to occur in contrastive contexts, have structures not well predicted by use of mutual exclusivity.

In contrast, for functional networks, those constructed by connecting words by shared relational, functional features, mutual exclusivity was a significant predictor of degree centrality when 2 or more overlapping features defined a connecting edge ($w = 2$). At this threshold, children who used mutual exclusivity at above-chance levels had semantic networks with higher degree centrality ($\rho = .34$, $p < .05$; Figure 3, bottom right). This same threshold is shown by Hills et al. (2009a) to best separate semantic categories in this set of words. Use of mutual exclusivity did not reliably predict clustering coefficient, but did show a positive trend, particularly at overlap threshold 3 ($\rho = .28$, $p = .1$; Figure 3, top right). This trend should be interpreted cautiously, however, as conceptual networks were quite sparse at $w = 3$, having at most 12 edges.
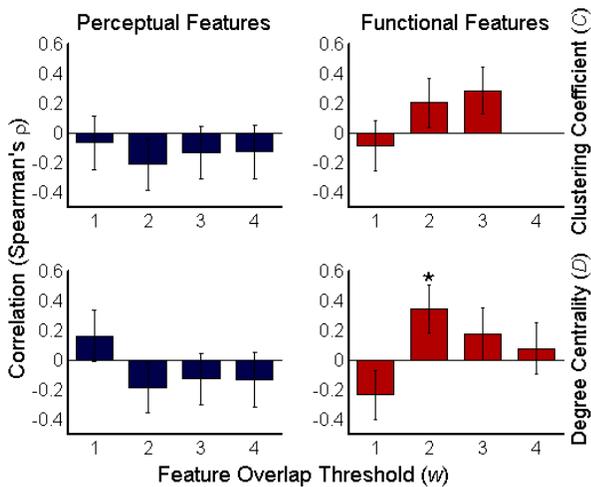
Figure 3: Correlation between network structure and mutual exclusivity performance. The top row shows correlations clustering coefficient, the bottom row for degree centrality. The left column shows perceptual features, the right shows functional. Individual bars shows correlations when particular thresholds (*w*) define the minimum number of overlapping semantic features required to connect two words. Error bars show +/-1 standard error as measured by 1000 samples of bootstrap resampling (Lunnenborg, 1985). Mutual exclusivity was significantly correlated with degree centrality in the functional networks and showed a non-significant trend towards correlation with correlation coefficient in these same networks. As predicted, ME was uncorrelated with perceptual network structure.

Thus, the semantic network structures of children who reliably exhibit mutual exclusivity are predictably different from those of children who do not. Even though these children know the same number of words, the words they know are different. Semantic networks of ME users are characterized by more hub-like structure, a consequence of the kind of word learning facilitated by exclusion. Importantly, these differences are likely to matter (Beckage, Smith, & Hills, 2011). Differences in connectivity structure lead to differences in network robustness, with the networks of mutual exclusivity structure perhaps protecting them against forgetting and aiding future learning (Albert, Jeong, Barabási). These results represent a first step in understanding how children's own learning mechanisms build the structure of their semantic networks.

General Discussion

While the words that children learn are, of course, a function of the linguistic input to which they are exposed (Brent & Siskind, 2001; Hills et al., 2009a), this link is likely to be moderated by children's own attentional and learning mechanisms (Hudson Kam & Newport, 2005; Smith, 2000). For instance, children learn to extend newly-learned object words to categories on the basis of particular feature dimensions. Normatively, children learn a bias to attend to shape, and this bias leads them to learn more categories organized by shape (Smith et al., 2002). However, children may learn a different bias, and consequently learn different words in the future (Colunga & Sims, 2011). We show that use of mutual exclusivity may play a similar role. Children who robustly use mutual exclusivity are likely to learn new words functionally related to words they already know. As atypical semantic network structure is related to slower language learning (Beckage, Smith, & Hills, 2011), these results point to a potential intervention for late-talking children. Learning to disambiguate the meanings of new words through exclusion could help late-talkers to catch up.

These results also lead to two further insights about mutual exclusivity and its role in vocabulary development. While mutual exclusivity is often thought to be critical to early word learning, its relationship to vocabulary size is unclear. For every study that finds a significant correlation between mutual exclusivity and vocabulary size (e.g. de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011; Mervis and Bertrand, 1994), another finds no correlation between the two (e.g. Halberda, 2003; Mather & Plunkett, 2009). These results help to shed light on this inconsistency by pointing out that the relationship between vocabulary development and mutual exclusivity may be found not in size but in structure. While we do not mean to argue that mutual exclusivity is required for rapid word learning, we do suggest that their relationship can be better understood by considering semantic network structure.

Finally, these results may shed light on the origins of mutual exclusivity itself. Thus far, we have argued that mutual exclusivity builds vocabulary structure. But

vocabulary structure may also build mutual exclusivity. One can think of mutual-exclusivity as an overhypothesis, a probabilistic rule about the general structure of word-object mappings derived from the structure of individual word-object mappings (Kemp, Perfors, & Tenenbaum, 2007; Mervis & Bertrand, 1994). For instance, mutual exclusivity may have its roots in an understanding that labels are often contrastive, pointing to differences between otherwise similar objects. If this is true, vocabularies that make this overhypothesis more probable should lead to stronger mutual exclusivity biases. Thus, one can think of the hub-like structures characteristic of ME users in our sample as not only arising from mutual exclusivity, but helping to construct it as well. Hub words, which are connected to many semantically-related neighbors, may play an important role in discovery of this higher-order regularity. Thus, mutual exclusivity may operate much like the shape bias: being both built from regularities in the structure of linguistic input, and helping children to discover further regularities (Smith, et al., 2002). A deep understanding of the connection between mutual exclusivity and vocabulary structure, then, will come from understanding a three part relationship: how ME contributes to structure, how structure contributes to ME, and language input contribute to both.

Acknowledgments

References

Albert, R., Jeong, R., & Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*, 378-382.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of verbal learning and verbal behavior*, *22*, 261-295.

Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B*, *364*, 3633-3647.

Beckage, N., Smith, L., & Hills, T. (2011). Small Worlds and Semantic Network Growth in Typical and Late Talkers. PloS ONE, 6, e19348.

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

Colunga, E., & Sims, C. E. (2011). Early talkers and late talkers know nouns that license different word learning biases. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2550-2555). Austin, TX: Cognitive Science Society.

de Marchena, A., Eigsti, I. M., Worek, A., Ono, K., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, *119*, 96-113.

Diesendruck, G., & Markson, L. (2001). Avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630–641.

Ebbinghaus, H. (1885/1962). Memory: A contribution to experimental psychology. New York: Dover.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59. Chicago: University of Chicago Press.

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, *1*, 215-239.

Golinkoff, R M., Hirsh-Pasek, K., Bailey, L., M. & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*, 99-108.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*, B23-B34.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009a). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, *112*, 381-396.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009b). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition?. *Psychological Science*, *20*, 729-739.

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Department of Sociology, Yale University, New Haven, CT.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151-195.

Lunnenborg, C. E. (1985). Estimating the correlation coefficient: The bootstrap approach. *Psychological Bulletin*, *98*, 209-215.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*, 307-321.

Markman, E., & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121-157.

Mather, E., & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*, *14*, 60-76.

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

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Development*, *65*, 1646-1662.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 92-134.

Smith, L. B. (2000). Learning to learn words: An associative crane. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51-80). New York: Oxford University Press.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object naming provides on-the-job training for attention. *Psychological Science*, *13*, 13-19.

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.