Title
Categorical Structure in Early Semantic Networks of Nouns

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

Categorical Structure in Early Semantic Networks of Nouns

Thomas Hills (thills@indiana.edu)
[1]Department of Psychological and Brain Sciences, 1101 E. Tenth Street
Bloomington, IN, 47405 USA

Mounir Maouene (mmaouen@ensat.ac.ma)
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

Despite what we know about children's ability to categorize, it is not clear to what extent information in the environment is capable of facilitating higher-order category knowledge, nor to what extent different kinds of object features play different kinds of roles. As a start we built a network of 130 early-learned nouns with 1394 perceptual and functional features as given by adult judgments. Then we analyzed the basic structural properties of the network. These revealed a small world structure and a high degree of feature overlap in local clusters. To identify the local clusters, we used a clique percolation algorithm to parse the network in terms of the statistical properties of feature overlap. This enabled us to identify clusters of items with a strong resemblance to common categories, such as animals, foods, and vehicles. Perceptual and functional features were found to play different roles in the categorization, with functional information being less redundant but more specific than perceptual information.

Keywords: early semantic network, clusters, perceptual and functional features, percolation algorithm, feature correlations.

Theories of human category structure are often based on feature-category correlations. Traditional theories of categorization posited necessary and sufficient defining features that determine category membership (reviewed in Murphy, 2002), but other important theories posit probabilistic feature correlations with category members typically being similar across clusters of correlated properties (Rosch et al., 1976). It is generally accepted that people learn category-feature correlations (McRae, Cree, Seidenberg & McNorman, 2005; Younger & Cohen, 1990) and there is supporting evidence both from developmental (Rakison & Poulin-Dubois, 2002) and category-specific deficit studies (for example, Caramazza & Shelton, 1998; Tyler, Moss, Durrant-Peatfield & Levy, 2002, but see Warrington & Shallice, 1984).

Computational studies suggest further that the latent structure available in a *system* of many categories with many feature-category correlations may be sufficient to define different superordinate kinds. In particular, Rogers and McClelland (2004) demonstrated how patterns of coherent co-variation across features could create superordinate categories. The idea of co-variation—of systems of correlations in overlapping feature patterns—provides one possible way to address the criticism that feature correlations are too-unconstrained to explain human category structure (e.g. Ahn, Kalish, Medin & Gelman, 1995).

Although previous studies have explored how feature-correlations structure specific categories, no prior studies have focused on how this structures the system of noun categories children learn. The purpose of the present study is to provide such a *description.* A descriptive study seems pre-requisite to the examination of any claims about what feature correlations—of themselves or in concert with other processes—can do by way of creating children's category knowledge. The present study specifically examined 130 nouns that are among the first nouns children learn and the structure of the features (derived from adult feature generation studies) associated with those categories. The analyses concentrate on perceptual and functional features--features of things that should be evident in even young children's experiences.

Perceptual and functional features are also of interest because of several disputes of the possibly different roles that the two kinds of features might play in category organization. Perceptual features are typically defined as static visual features such as color, shape, and part structure (e.g., having legs or not); functional features typically encompass roles (e.g., used to drink from), behaviors (breathes or flies), and transient properties (e.g., can be opened). There have been suggestions that different kinds of categories differ in the relative importance of these two kinds of features with perceptual features perhaps more important for animals and functional features more important for artifacts, although there is considerable dispute (see De Renzi & Lucchelli, 1994; Komatsu 1992). In the literature on cognitive development, the debate centers around the relative importance of the two classes of features, with some suggesting a developmental trend from more perceptual to more functional features in category organization and others suggesting that functional features trump perceptual features even early in category development. (see, e.g., Mandler, 1992; Quinn and Eimas, 1996; Sheya and Smith, 2007).

In summary, the goal of this study is a *description* of the coherence of the co-variation of the feature category correlations that characterize the noun categories children typically learn early, prior to the age of 3 years. Such a description should enable fine-grained behavioral predictions about, for example, children's generalizations from one category to another and about the early organization of nascent superordinate categories. Such a description should also be relevant to understanding the kinds of learning mechanisms that might exploit the available structure. To these ends, we examine the graph theoretic properties of the system of pair-wise relations among nouns as indicated by the numbers of shared features that connect them. The choice of this approach is meant to be atheoretic; that is, we seek simply to describe the structure potentially available to some learning mechanism, not to show that any such structure or generalization from that structure can be learned. In the analyses, we will compare the graph theoretic structure in the feature network of these early-learned nouns to randomly connected networks, to better understand how much and what kind of structure is available to young learners.

The Categories and Their Features

Noun categories. The nouns were selected from the Bates-MacArthur Communicative Developmental Inventory (Fenson, Dale, Reznick & Bates, 1994), Toddler version. This inventory contains at least 50% of children's productive vocabulary by 30 months. Our study used the 130 common nouns from this inventory for which there were available feature norms from McRae et al., (2005). These 130 nouns over-represent (with respect to the inventory as a whole) animals (33 nouns, 25% of the subset versus 15% of whole inventory) and under-represent food (17 nouns, 13% of the subset versus 23% of the whole inventory). Nonetheless, the sample includes a broad array of nouns across several different superordinate categories.

The features. The features were taken from the feature norms reported by MacRae et al. (2005). That study collected feature norm 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. Cree and McRae (2003) classified the generated features as perceptual (e.g., is red, has wheels), functional (e.g., used to eat food, gives rides), encyclopedic (e.g., was imported to Europe from the New World) and taxonomic (e.g., is an animal). We used only the perceptual and functional features because only these types of features are likely to be directly experienced by young children. Note that superordinate names, the likely real-world correlates of the taxonomic features, are not typically known by children younger than age three years. One limitation of this approach (well-recognized in the feature generation literature) is that the features generated by adults exclude any pervasive and/or not easily labeled properties (e.g. kinds of shapes or ubiquitous behaviors such as "breathes"). However, adults do consistently list features that are characteristic of things. Thus, although the features generated by adults are likely to be incomplete with respect to those available to learners, the structure discernible from such imperfect data is nonetheless likely to be informative about the coherent co-variations of features among early noun categories.

The Structure of Early Noun-feature Correlations

The Full Network. A graph is a collection of nodes and a collection of edges that connect pairs of nodes. In the following analyses, the edges may be defined in terms of differing numbers of shared features: for example, when w (the feature threshold to define an edge) is 1, nouns are connected by an edge if they share at least one feature and when w is 3, nouns are connected by an edge if they share at least 3 features. In all the analyses, the threshold w was varied between 1 and 4. These different criteria for defining edges (and the connectedness of any two nouns) yield a series of networks.

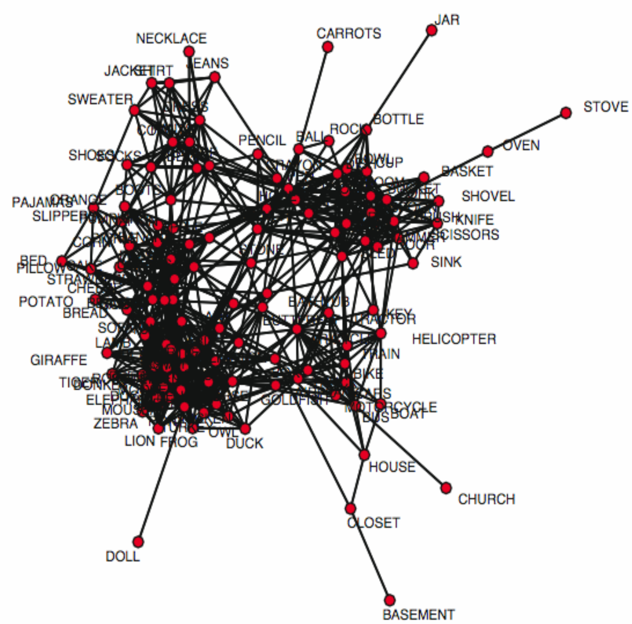

Figure 1: The noun graph. Edges represent ($w = 2$) two or more shared features. Isolated nodes are not shown.

Small-world structure. Our first question was whether or not our feature-based noun network contained sufficient local structure to warrant further investigation. A common method used to categorize the presence of local network structure is small-world analyses. Watts and Strogatz (1998) showed that many real world networks contain small-world structure. Additionally Steyvers and Tenenbaum (2005) showed that several representations for adult semantic networks contain small-world structure as well. Watts and Strogatz (1998) found that the small-world structure in the real world networks had a mean clustering coefficient much larger than the one found in a random network of the same number of nodes, edges, and average shortest path length. The clustering coefficient of a node was calculated by determining what proportion of a node's closest neighbors (nodes connected by an edge) were also connected to each other by an edge. To get the clustering coefficient of the whole network, the coefficient was averaged across all nodes. Consequently, when the average shortest path length of a network is low and the mean coefficient cluster is high relative to the appropriate random network, the network is said to contain *small-world structure*.

The feature-based network we present here has small-world structure, which increases with increasing *w*. As *w* increases from 1 to 4, the clustering coefficient increases from 0.55 to 0.6, while the mean clustering coefficient of 500 random networks of the same size and density goes from 0.29 to 0.02. The mean shortest path length varies from 0 to 5 in both the observed and random networks, but does so in a non-systematic way. In fact the mean shortest path length decreases faster as a function of the feature threshold *w* for the random network as compared to the feature-based network. What this means is that the feature-based network has a more robust small world structure than the random network or that it becomes *more* locally structured as the feature requirement for edges increases. Thus, comparing the mean coefficient cluster and the mean shortest path length of both random networks and the feature-based network helped us determine that there was sufficient local structure to warrant an investigation into the identity of the underlying clusters.

Cluster Analyses. Because we are interested in how the coherent co-variations of features structure early noun categories, we needed to employ a clustering technique which did not force strict partition of noun clusters. In fact this might be a more principled way to explore the semantic knowledge that structures the flexible system of early noun categories. For example, CHICKEN may belong in a category with *birds* or other *animals*, but it may also share certain features with members of the *food* category. Given that particular goal, we used the clique percolation method introduced by Palla, Derenyi, Farkas, and Vicsek (2005) that does precisely that.

The clique percolation method identifies groups of nodes of size k that are well-connected with one another. It does this by identifying the presence of *k-cliques*, which are sets of *k* nodes that are all connected with one another (maximal complete subgraphs). Two *k-cliques* are adjacent if they share *k*–1 vertices and two *k-cliques* are *k-clique-connected* if they are connected by a sequence of adjacent *k-cliques*. A *k-clique percolation cluster* is the union of all *k-cliques* that are *k-clique-connected* to a particular *k*-clique. For a given value of *k*, this method identifies all *k-clique percolation clusters*, which are groups of objects that are sufficiently well connected to create clusters of local structure in the network.

Besides allowing for overlapping categories and using local edge densities to identify communities of nouns, the clique percolation method also allows for a principled approach to identifying the cut-off threshold for a given network that yields the most structural information (see Palla, Barabasi, & Vicsek, 2007). We accomplish this by increasing the value of *k* for each cut-off threshold, *w*, until the second largest component is larger than half the size of the largest component. This is because, for low values of *k*, most nodes tend to be connected in one large clique percolation cluster. However, as *k* is increased, the clique percolation clusters separate as the method focuses in on narrow regions of high connectivity. After adjusting *k* upwards for each cut-off threshold, we then identify the corresponding *w* and *k* that have the largest number of percolation clusters. For the noun-feature network, the *k* and *w* values that yield the most clusters are 3 and 3, respectively. This yields a conservative estimate for category membership, because only nouns with enough local information to be included in a clique of size $k = 3$ will be included in the output. Nouns lacking this connectedness are not assigned to any cluster.

The largest number of clusters identified was 10 for $k = 3$ and $w = 3$ (see Appendix for clusters). These clusters represent potential category structure and are somewhat consistent with our adult expectations, at least in terms of what they include. For example, there are categories approximating what adults may define as food, vehicles, non-flying animals, birds, clothing, furniture, and several categories for artifacts. Comparing the adult-reported taxonomic membership (see superscripts in the Appendix) with the percolation clusters finds significant parallels. Categories that are most clearly consistent with adult taxonomic categories are those of animals, fruits, birds, clothing, furniture, and vehicles. In all cases, some members lie outside the taxonomic assignment of McRae et al. (2005), some of which would likely be classified as category members by most adults, while some of which would not. For example, AIRPLANE, BIKE, and TRACTOR do not have 'vehicle' assignment according to McRae et al. (2005), but few adults would probably argue that they are not vehicles. Of perhaps equal interest is that the feature clusters pick up *ad hoc* categories (Barsalou, 1983) such as a category of things for cutting, a soft-white things category, and a category of things you might need to rest and relax. The categories—defined only by the connectedness of the nouns through shared features—

constitute hypotheses about young children's category knowledge that can be tested in future behavioral studies.

In sum, the results suggest two main results. 1) Common categories known by young children present highly organized local structure in terms of shared features. This local structure may constitute children's early higher order knowledge about different categorical kinds. 2) It also shows that things belong to different partitions and that these partitions are organized along coherent co-variations of overlapping and co-occurring features.

The Functional and Perceptual Networks

To address the structural roles of perceptual and functional features in the noun-feature network, we used the perceptual and functional classifications assigned by McRae et al. (2005). The noun-feature network contains 403 functional features (3.1 per noun) and 991 perceptual features (7.6 per noun). The most common functional features (in terms of the number of nouns with which they are associated) are: *is edible* (20), *used for transportation* (11), *worn for warmth* (8), *hunted by people* (6), *used by children* (6), *used for holding things* (6). The most common perceptual features are: *made of metal* (24), *can be different colors* (22), *has 4 legs* (22), *is large* (21), *is small* (21)).

Functional and perceptual networks were generated by creating networks that only used functional or only used perceptual features. These networks therefore represent different kinds of semantic information that could structure the system of early noun categories in different ways. Figures 2 and 3 present the functional and perceptual networks, with $w = 1$ for the functional network and $w = 2$ for the perceptual network. At $w = 1$ the perceptual network is one large component with no visible local structure.

As is apparent from Figures 2 and 3 (and the associated threshold w), the perceptual network is far denser than the functional network. On average, a node in the perceptual network at the lowest cut-off threshold is connected to 27% of the other nodes; the average node at the same cut-off threshold in the functional network is only connected to 5% of the other nodes. Perceptual information is therefore more noisy and, in a category defining sense, less discriminating than functional information.

The number of isolates is much higher for the functional network than for the perceptual network. At a cut-off threshold of $w = 2$, more than half of the nodes in the functional network are unconnected to any other node. At the same cut-off threshold, only 10 nodes in the perceptual network are isolates. This indicates that perceptual noun relationships tend to be more redundant than functional relationships. Edge relationships in the functional network are predominantly based on a single shared feature.

Both networks have small-world structure. With w ranging from 1 to 4, the functional network clustering coefficients go from 0.88 to 1. For the same w range, the perceptual network clustering coefficients go from 0.54 to 0.62. However, at $w = 2$, the number of isolated nodes in the functional network is 81, but only 10 for the perceptual network. The difference between the observed clustering coefficients and that for a random network of similar density is higher for the functional network than for the perceptual network, which is consistent with what we can visually observe in Figures 2 and 3: the functional network has more local structure than the perceptual network. However, even the slightest increase in the cut-off threshold reduces the functional network to a large number of isolates. Meanwhile, the perceptual network maintains small-world structure and involves the majority of the nodes in this structure even if the requirement for noun-pair relatedness is three or more perceptual features. There is a very clear trade-off here. Perceptual information, partially because of its abundance, is more redundant and can provide more robust information about category inclusion, but this information is not as discriminating of different categories as is functional information. A single functional relationship is sufficient to define all category members that are, for example, USED FOR TRANSPORTION; no single perceptual feature contains that information.

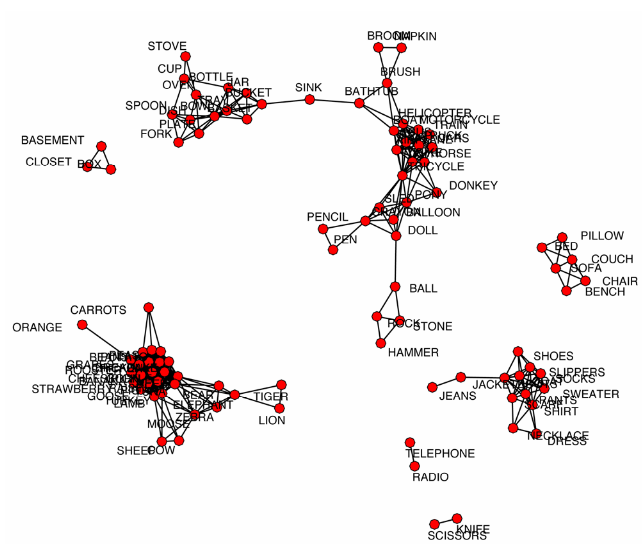

Figure 2: The functional feature network ($w = 1$). Isolated nodes are not shown.

Cluster Analyses. What local structure is present in the perceptual and functional networks? The functional network provides the most number of clusters (11) when $k = 3$ and $w = 1$; for the perceptual network, the most clusters (9) are separated out when $k = 5$ and $w = 2$. This is consistent with the graph theoretic data showing that the functional network has fewer isolates and greater local structure at its lowest cut-off threshold, while the perceptual network loses only a few nodes to isolates but gains substantial local structure—compared with a random network of the same density—by increasing w to 2.

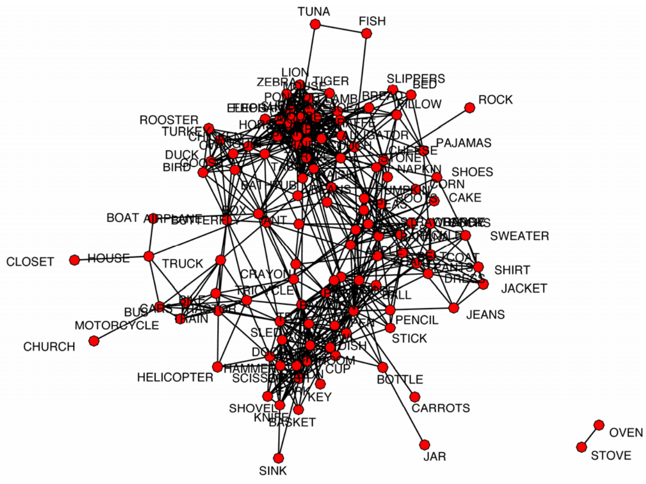

Figure 3: The perceptual feature network ($w = 2$). Isolated nodes are not shown.

For the sake of brevity, we do not present the clique percolation clusters here, but instead describe their relevant properties. A close look at the different kinds of clusters present in the two networks reveals some interesting comparisons. First, the clusters in the functional category are generally smaller, more conservative—there are fewer odd members in any category—and tend to represent what adults might call traditional categories (e.g., food, clothes, vehicles). The perceptual based clusters are less conservative—containing numerous odd members—and are more representative of *ad hoc* categories. For example, to our best approximation, two of the perceptual clusters might represent LONG THIN THINGS and THINGS THAT CAN FLY. As above, comparing the clique percolation clusters with the taxonomic information reported by adults in the McRae study is not particularly helpful. For example, TUNA is not labeled as *food* in the McRae study, but it is clearly recognized as such in the functional clique percolation clusters. As well, the functional clusters recognize categories of THINGS TO RIDE IN OR ON, for which there is no taxonomic category in McRae. Nonetheless, the functional categories contain only food items in food clusters, whereas the perceptual clusters contain PEN and HOSE in with other fruits and vegetables, and a larger mostly animal category, contains furniture (having 4 legs) and BOX (is brown and can be large or small, like MOOSE and SQUIRREL). Overall, functional categories tend to be smaller (underestimating category membership) and less sullied by near-members, whereas perceptual categories are larger and over-estimate category membership. Again, these data provide category specific hypotheses through which to test different feature roles in category membership.

Remembering that our most discriminating perceptual clustering requires two features to create a relationship, while the functional clusters require only one, one cognitive interpretation is that categories based on perceptual features should require more information to make comparisons (including possibly more feature overlap and more exemplars), whereas functional categories can use very limited information, but are constrained by the limited availability of that information. Perhaps the right conclusion will be that perceptual and functional features play important but different roles in developing category knowledge.

General Discussion

The capacity to create flexible categories from feature correlations and overlap is a powerful tool for predicting properties about the world. By taking a subset of nouns that many children know at 30 months and combining these with features reported to be characteristic of these things, we were able to construct a network that represents a cognitive hypothesis about how information is structured in early semantic networks. Analyses of this network revealed that it had small-world structure and that the local structure was consistent with categories that are largely familiar as *ad hoc* categories of practical utility.

Further analyses revealed that the functional and perceptual features that made up the network played different kinds of roles in structuring the network. Functional features tended to play a more conservative role, in most cases only including more traditional members of standard adult categories. Perceptual features were more redundant, with multiple features defining inclusions in a specific cluster, but were also less discriminating, and capable of producing clusters that adults do not typically identify as categories (e.g., long-thin-things) but which young children might.

This is a first encouraging step in understanding how systems of category-feature correlations and overlap might constitute category knowledge and a step that leads to specific testable hypotheses. However the description given here might well be distorted by the very limits of the tools we used. There are a number of limitations in the use of the adult-generated features. First, as McRae et al. (2005) point out, they are linguistically based. Thus, potentially highly relevant properties that are hard to describe in words (the shape of cows versus horses, or the relative sizes of things) may be missed. Second, generated features are likely to be distorted by cognitive biases and for example, emphasize distinguishing features at the basic level (e.g., has webbed feet) over common features at the superordinate level (e.g., breathes). However, given that the feature norms do represent an under-estimate of available features and are unlikely to focus on shared categorical features, the presence of categorical structure in the data is strong evidence that feature-correlations in the environment are sufficient to produce categorical inferences in children.

Acknowledgments

This work was supported by the NICHHD, grant RO1 HD028675.

References

Ahn, W. K., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition,* 54, 299-352.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition,* 11(3), 271-227.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience,* 10, 1-34.

Cree, G. S., & 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.

De Renzi, E., & Lucchelli, F. (1994). Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex,* 30, 3-25.

Fenson, L., Dale, P. S., Reznick, J., & Bates, E. (1993). Variability in early communicative development. *Monographs of the Society for Research in Child Development,* 59(5), v-173.

Komatsu, L. (1992). Recent views of conceptual structure. *Psychological Bulletin,* 112(3), 500-526.

Mandler, J. M. (1992). How to build a baby II: conceptual primitives. *Psychological Review* 99(4), 587-604.

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.

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge MA: MIT Press.

Palla, G., Barabasi, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature,* 446(5), 664-667.

Palla, G., Derenyi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature,* 435(9), 814-818.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual organization and categorization in young infants. In C. Rovee-Colier & L. P. Lipsitt (Eds.), *Advances in infancy research*. Stanford CT: Ablex.

Rakison, D. H., & Poulin-Dubois, D. (2002). You'll go this way and I'll go that way. Developmental changes in infants' detection of correlations among static and dynamic features in motion events. *Child Development,* 73, 682-699.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge MA: MIT Press.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology,* 8(3), 382-439.

Sheya, A., & Smith, L. B. (2007). Perceptual features and the development of conceptual knowledge. *Journal of Cognition and Development,* 7(4), 455-476.

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.

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. *Brain and Language* 75(2), 195-231.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain,* 107, 829-854.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature,* 393, 440-442.

Younger, B. A., & Cohen, L. B. (1990). Infants' detection of correlation among feature categories. *Child Development,* 61(3), 614-620.

Appendix

Complete Noun-Feature Network Clique Percolation Clusters

Superscripts provide the taxonomic category (if present) in the McRae et al. (2005) norms. When the superscript is a single letter (or letters), it is defined at its first appearance as the most recent superscripted word starting with the same letter (or letters). For example, f is FOOD.

0: bread[food] cheese[f] lamb[animal]
1: balloon bench bowl broom brush bucket comb cup[dish] dish fork[utensil] hose[tube] pen plate[dish] sled spoon[utensil] tray
2: door hammer[tool] shovel[t] spoon[u]
3: fork[u] knife[u] scissors[t]
4: bear[mammal,a] cat[m,a,pet] couch[furniture] cow[m,a] deer[m,a] dog[m,a,p,carnivore] donkey[m,a] elephant[m,a] frog[a, amphibian] horse[a] lamb[a] lion[m,a,carnivore,cat,predator] moose[m,a,herbivore] mouse[a,rodent] pig[m,a] pony[a,horse] sheep[m,a] squirrel[m,a] tiger[a,predator,cat,carnivore] turtle[a,pet,reptile,amphibian] zebra[m,a,horse]
5: boots[clothing] coat[cl] dress[cl] jacket[cl] jeans[cl] pants[cl] scarf[cl] shirt[cl] shoes[cl] slippers[cl] socks[cl] sweater[cl]
6: apple[fruit] banana[fr] beans[vegetable] cake[dessert] cheese[food] grape[fr] orange[fr] peas[vegetable] pickle[vegetable, food] raisin[fr] strawberry[fr] tuna[fish]
7: bear[mammal,a] blackbird[bird,a] chicken[b,a] duck [b,a] goose [b,a] owl [b,a] penguin [b,a] rooster [b,a] turkey [b,a]
8: bed[furniture] chair[fu] couch[fu] pillow slipper[cl] sofa[fu]
9: airplane bike bus[vehicle] car[v] horse[a] motorcycle[v] tractor[machine] train tricycle[v,bike] truck[v]