Published in final edited form as:

J Exp Psychol Gen. 2019 April ; 148(4): 763–782. doi:10.1037/xge0000596.

Modeling Early Lexico-Semantic Network Development: Perceptual Features Matter Most

Ryan Peters and Arielle Borovsky

Department of Speech, Hearing, and Language Sciences, Purdue University, West Lafayette, IN

Abstract

What aspects of word meaning are important in early word learning and lexico-semantic network development? Adult lexico-semantic systems flexibly encode multiple types of semantic features, including: functional, perceptual, taxonomic, and encyclopedic. However, various theoretical accounts of lexical development differ on whether and how these semantic properties of word meanings are initially encoded into young children's emerging lexico-semantic networks. While some accounts highlight the importance of early perceptual vs. conceptual properties, others posit that thematic or functional aspects of word meaning are primary relative to taxonomic knowledge. We seek to shed light on these debates with two modeling studies that explore patterns in early word learning using a large database of early vocabulary in 5450 children, and a newly developed set of semantic features of early-acquired nouns. In study 1, we ask whether semantic properties of early-acquired words relate to order in which these words are typically learned and study 2 models normative lexico-semantic noun-feature network development compared to random network growth. Both studies provide converging evidence that perceptual properties of word meanings play a key role in early word learning and lexico-semantic network development. The findings lend support to theoretical accounts of language learning that highlight the importance of the child's perceptual experience.

Keywords

Lexico-semantic development; Semantic networks; Word learning; Language development; Semantic Features

1. Introduction

What kinds of word meanings do young children know and how does this lexico-semantic knowledge change across development? One way to classify sub-components of word (e.g., CAT) meanings is to use semantic features (e.g., *has fur*, *has four legs*, *is a pet*, *meows*, etc.; McRae, De Sa, & Seidenberg, 1997; Rogers & McClelland, 2004). Researchers can create semantic features (e.g., Rumelhart, 1990) or empirically derive them by collecting semantic feature production norms (e.g., McRae, Cree, Seidenberg, & McNorgan, 2005). Empirically

derived semantic features explain variance in numerous empirical findings in adult psycholinguistic processing, including: semantic priming (McRae et al., 1997; Vigliocco, Vinson, Lewis, & Garret, 2004), categorization (Hampton, 1997; Smith, Shoben, & Rips, 1974), conceptual structure (Hills et al., 2009a; Medin, 1989), semantic richness (Pexman et al., 2008; Rabovsky, Sommer, & Abdel Rahman, 2011), and word/concept learning (Engelthaler & Hills, 2017; Hills et al., 2009b; Hills, Maouene, Riordan, & Smith, 2010; Landau, Smith, & Jones, 1988; Rogers & McClelland, 2004; Yurovsky, Bion, Smith, & Fernald, 2012). However, preexisting sets of semantic feature norms miss many early-learned words, limiting the adoption of the featural approach to the study of lexical development. Here, we use a newly-developed set of feature norms for early-acquired nouns (Peters, McRae & Borovsky, in prep) to measure associations between the features of early-acquired words and Age of Acquisition (AoA), and explore how these features contribute to early lexico-semantic network structure. Before describing our approach, we review contrasting theoretical accounts of early lexico-semantic learning and representation, and recent work that re-frames the issue in terms of lexico-semantic noun-feature networks to circumvent limitations of earlier methodologies.

Numerous theories of adult lexico-semantic representation posit that word meaning can be characterized in terms of semantic features (e.g. Collins & Quillian, 1969; Rogers & McClelland, 2004; Smith & Medin, 1981). These featural approaches highlight several possible types of features, including: functional, perceptual, taxonomic, and encyclopedic features (Cree & McRae, 2003). *Functional features* refer to functional or motor information regarding the ways in which people interact with objects (e.g., *used by children, used for sports*, etc.), and are closely related to thematic, associative, or role-role relations that characterize the co-occurrence of objects in different situations. *Perceptual features* refer to information that is accessible through a primary sensory channel (e.g., *has 4 legs, barks*, etc.). *Taxonomic features* describe relational positions in a taxonomic, or categorical, hierarchy (e.g., *an animal, is a food*, etc.). Finally, *encyclopedic features* refer to facts that don't fall into any of the other three categories (e.g., *has businesses, is dangerous*, etc.).

While adult lexico-semantic systems flexibly encode all the above, researchers still debate when young children know different types of semantic features. The *thematic-to-taxonomic shift* (e.g. Inhelder & Piaget, 1964; Smiley & Brown, 1979) posits that thematic (a.k.a. associative or functional) relations initially drive semantic processes in children, and only later in development do children prioritize other kinds of relations. For example, in a picture triad preference task, when presented with NEEDLE and asked if it goes better with either THREAD or PIN, kindergarteners are more likely to choose the THREAD – which shares an associative relation by participating together in sewing events – while older children and adults are more likely to choose PIN; which, like NEEDLE, is also a sharp sewing tool (Smiley & Brown, 1979). This developmental pattern has much in common with other proposals, such as theory-theory and core-knowledge approaches (e.g. Carey, 1985; Gelman, 1990; Gopnik & Meltzoff, 1997; Keil, 1989; Murphy & Medin, 1985; Spelke, 1994), which propose that early in semantic development object features that are not directly available in the perceptual input – such as causal relations that require some degree of inferential knowledge to parse – play a central role in categorization and word learning. Such features

are often described as conceptual features, but some researchers liken them to the functional features described above (e.g., Hills et al., 2009b).

In contrast, proponents of a *perceptual-to-conceptual shift* (Quinn & Eimas, 1997) suggest categorization and word learning initially depend on the perceptual features of objects, and children only begin to represent more complex/abstract features and relations after the lexico-semantic system develops sufficiently. For example, an infant may first recognize perceptual features such as eyes, appendages, body shape and species-specific sounds when they begin to learn the concept ANIMAL. Then, only later, and at times with the aid of language, do they solidify the representation with more abstract features such as biological functions and taxonomic relations to other concepts (Quinn & Eimas, 2000). This idea is shared by theories based on association learning mechanisms – in which lexico-semantic development is proposed to be a “process of continuous representational enrichment grounded in a perceptual system that is sufficiently sensitive to allow infants to form categories that cohere because of perceptible similarity relations” (Rakison et al., 2008, p. 13; Hollich et al., 2000; Quinn & Eimas, 1997, 2000; Samuelson & Smith, 1998, 1999; Smith, 2000; Smith & Heise, 1992; Smith, Jones, & Landau, 1996; Smith, Jones, Landau, Geshkoff-Stowe, & Samuelson, 2002). Supporting work demonstrates that infants learn perceptual regularities across objects (Quinn & Eimas, 1996; Rakison, 2003; Younger, 1990) but initially struggle to recognize and use conceptual regularities (e.g. Madole, Oakes, & Cohen, 1993; Rakison, 2005; Rakison & Cohen, 1999; Sheya & Smith, 2006), and that iconicity (Caselli & Pyers, 2017; Perry, Perlman, & Lupyan, 2015; Perry et al., 2017) and perceptual salience (Pruden et al., 2006) facilitate early vocabulary learning.

However, these theoretical proposals are underdetermined due to limitations that constrain the scope of past research. First, infants cannot complete many tasks used to assess lexico-semantic processing in older children and adults. This limitation has forced researchers to create methods that do not require infants to intentionally engage in a task, such as: familiarization-novelty preference procedures (e.g., Quinn & Eimas 1996), sequential-touching tasks (e.g., Gopnik & Meltzoff, 1987; Rakison & Cohen, 1999), and generalized imitation tasks (e.g., Mandler & McDonough, 1996). These techniques have advanced our understanding of infant categorization and word learning. However, they only indirectly assess lexico-semantic processing. Consequently, different techniques produce different and sometimes-conflicting results – which in turn support divergent theoretical views (Oakes & Rakison, 2003). Recent advances in infant-appropriate semantic priming methods (e.g., Arias-Trejo & Plunket, 2009; 2013; Friedrich & Friedrici, 2006) – such as the looking-while-listening paradigm (e.g., Borovsky, Ellis, Evan & Elman, 2016) and other auditory semantic priming tasks (e.g., Rämä, Sirri, & Serres, 2013; Willits, Wojcik, Seidenberg, & Saffran, 2013) – promise new insights into these questions. However, to date, these empirical studies have been limited to exploring semantic relations that are either quite general (comparing semantically “related” vs. “unrelated” conditions) or can be clearly differentiated (“taxonomic” or “associative”). However, as discussed above, conceptual relations often have a rich, multi-featured substructure (e.g., CATDOG share perceptual, taxonomic, and encyclopedic features). Another limitation is that the structure of semantic connections between words have generally been described via metrics that are overly simplistic and/or limited in grain size. For example, prior research describing the semantic

neighborhoods of words have typically only looked at the number of neighbors (e.g., Storkel & Adlof, 2009), ignoring possible complex structure between neighbors and the possible role of large-scale structure beyond the neighborhood. These aspects of structure are important given modeling evidence that the broader semantic web surrounding a novel word plays a crucial role in learning (e.g., Landauer & Dumais, 1997; McMurray, Horst, & Samuelson, 2012). Below we discuss a promising way forward that circumvents these limitations: modeling early lexico-semantic development as growing noun-feature networks – an approach that builds on advances in the application of graph-theoretic network modeling techniques to lexico-semantic representations (e.g. Hills et al., 2009a; 2009b; Steyvers & Tenenbaum, 2005).

1.1. Using graph-theoretic noun-feature networks to explore early lexico-semantic development

Noun-feature networks can be visualized as a conglomeration of *nodes* and *links* - where nodes represent noun concepts and links represent shared features according to a set of feature production norms (e.g. McRae, Cree, Seidenberg, & McNorgan, 2005). For example, figure 1 shows SPOON and five concepts that it shares features with.

To collect feature production norm datasets, participants are presented with a list of concepts and asked to provide properties that come to mind for each concept. This procedure provides insight into conceptual representations “not because they yield a literal record of semantic representation (i.e., they do not provide a verbatim read-out), but rather because such representations are used systematically by participants when generating features (Barsalou, 2003)” (McRae et al., 2005 p.549). In other words, feature norms are not definitive psychological representations, but, as discussed above, they do capture aspects of word meaning that contribute to a range of behavioral phenomena that reflect lexico-semantic representation. While other methods to define semantic links between concepts are possible – such as corpus-based co-occurrence statistics (e.g., Beckage, Smith, & Hills, 2011; Hills 2013) or adult association norms (e.g., Steyvers & Tenenbaum, 2005) – feature-based links impart unique advantages in the study of early lexico-semantic development. Namely, noun-feature networks enable simultaneous characterization of *which* concepts are connected and *how* they are connected – illustrated in figure 1 by the different types of features forming the substructure of links between SPOON and its neighbors. Using this capability, noun-feature networks can clarify how different types of semantic features influence early lexico-semantic network structure, which in turn may afford interesting insights into word learning processes.

Hills and colleagues (2009b) compared early noun-feature networks with links formed via either perceptual or conceptual (i.e. functional) features to clarify how features of each type influence early lexico-semantic network structure. Both kinds of noun-feature networks contained clusters of words that resemble adult superordinate taxonomic designations (e.g., toys, animals, clothes, food, etc.). However, perceptual networks contained more links and had a higher percentage of links made up of multiple correlated features. Feature correlations – for example, how things that *have wings* are also likely to *have beaks* and *have feathers* – influence adult performance in semantic tasks (McRae et al., 1997; McRae, Cree,

Westmacott, & De Sa, 1999). Thus, such structural patterns in early noun-feature networks might relate to the ease with which different words are learned. Engelthaler and Hills (2017) explored this issue by looking at relations between semantic distinctiveness of early-learned words and adult-rated AoA. They found a significant negative relation between semantic distinctiveness and AoA that was largely driven by perceptual features. This result supports proposals that perceptual features play a particularly important role in early lexico-semantic development.

The above research demonstrates the utility of noun-feature networks to advance understanding of lexico-semantic development. However, several limitations in this work motivated the current project. First, prior analyses used feature norm datasets (e.g. McRae et al., 2005) that miss many early-learned words. This limited Hills and colleagues (2009a) to exploring networks using only 130 of the 359 early-learned nouns in the MacArthur Bates Communicative Developmental Inventory: Words and Sentences (MBCDI; Fenson et al., 1996), and led Engelthaler and Hills (2017) to use words that young children are unlikely to know in conjunction with adult generated AoA values. Second, although the reviewed analyses have explored both large-scale (e.g., clustering) and more local measures of network structure (e.g., distinctiveness), they are yet to explore whether these findings may be driven by more granular measures that *only* consider the immediately local information at the word- or link-level. Importantly, the structural locus of these effects (e.g. word-level, neighborhood-level, and lexicon-level) carries implications for models of lexico-semantic development. For example, evidence suggests that large-scale structural patterns facilitate long-range spread of lexical activation, supporting fluent processing and word retrieval (Griffiths, Steyvers, & Firl, 2007; Steyvers & Tenenbaum, 2005). However, evidence also suggests more local-level structure can facilitate processing and learning among words in a smaller-scale, shared semantic neighborhood (Borovsky & Peters, under review), and simple measures of the features that comprise a word may inform us about how the semantic properties of the word itself drives early learning.

Similar or dissimilar semantic properties can potentially drive larger and smaller-scale structure in several ways. One possibility is that words with more perceptual features are learned earlier than those with fewer perceptual features. This word-level pattern could then, in turn, additively relate to larger-scale network structure via perceptual features that are shared between word meanings. Here, relations between large-scale structure and the order of lexical acquisition could be said to arise out of the ease of learning individual words, rather than network development per se. This pattern may be analogous to attentional accounts of early word learning that point to the importance of perceptual correlations among objects to drive word learning and categorial structure (e.g. Smith et al. 2002, Smith & Heise, 1992). Alternatively, word-level feature preference might not drive large-scale structure in the network. In the word-level perceptual feature example above, perceptual features that comprise earlier-acquired words might NOT overlap with later-acquired words – and thereby would not drive higher-level lexico-semantic connections between words. This pattern is consistent with accounts highlighting the importance of semantic distinctiveness in early word learning (e.g., Engelthaler & Hills, 2017). Finally, it is also possible that word-level perceptual features do not relate to early word learning, but higher-level structural characteristics do. In other words, nonlinear relations between lower and higher-level

structure may result in emergent large-scale characteristics that relate to early word learning. As higher-level structure in perceptual feature networks resembles adult superordinate taxonomic designations (Hills et al., 2009b), in such a case one might also expect other kinds of features (e.g. functional, taxonomic), though not as frequent, to share more overlap and instead explain large-scale network connectivity. This potential outcome would be consistent with conceptual accounts of early representation that highlight the roles of associative and taxonomic links in early lexico-semantic structure (e.g., Unger, Fisher, Nugent, Ventura, MacLellan, 2016). Thus, comparing structure at multiple levels of granularity, ranging from the single-word level, to large-scale lexico-semantic network structure, can advance our understanding of mechanisms that drive lexical development.

To address whether and how differing semantic properties contribute to word learning and how these properties interact at different levels of lexical structure, we carry out two modeling studies, each with contrasting strengths and weaknesses. In study 1 we use multiple linear regression to develop a single statistical model that measures linear relations between the words' AoA and semantic feature characteristics at three levels: local (word)-, neighborhood-, and lexicon-level, while controlling for other potentially confounding variables such as word frequency. However, this regression approach has a key limitation: following standard practice (e.g., Engelthaler & Hills, 2017; Hills et al., 2009a; Steyvers & Tenenbaum, 2005), we calculate network metrics for neighborhood- and lexicon-levels of structure using the full lexicon networks rather than the networks that would exist at the point in development when a word is acquired. In contrast, in study 2, we use a novel approach that overcomes the limitation of the regression approach used in study 1 to measure how different semantic features drive global nonlinear developmental trajectories in early lexico-semantic network development. However, this approach, in turn, is limited in that the developmental pattern for each measure of network structure must be investigated individually. Another difference between the methods of studies 1 and 2 that deserves further clarification is the unit of analysis: in study 1 the unit of analysis is the word, while in study 2 it is the entire lexicon. Thus, across studies 1 and 2, we use methods that contrast in terms of both strengths and weaknesses, and in focus on local versus global measures of lexico-semantic network structure, to look for converging evidence on whether and how differing semantic properties contribute to early word learning.

2. Study 1: Regression analyses of Age of Acquisition using local-, neighborhood-, and lexicon-level measures of lexico-semantic network structure

To explore how different aspects of word meaning drive early word acquisition, we begin by measuring linear relations between several dimensions of word meaning and Age of Acquisition (AoA) at several levels of lexico-semantic structure. Our dimensions of meaning include those noted in our earlier review: perceptual, functional, taxonomic and encyclopedic. We measure relations between these featural components of lexical meaning and AoA at several levels of structure: local (word)-, neighborhood-, and lexicon-level. At the local (word)-level, we assess the featural makeup of a word defined by the Number of Features (NoF) of each type. This measure is generally taken to reflect the semantic richness

of a concept (Pexman, Lupker, & Hino, 2002; Pexman et al., 2008). At the neighborhood-level, we focus on metrics that describe the structural characteristics of the immediate neighborhood of a word: Degree and Clustering Coefficient. Finally, at the lexicon-level, we focus on a metric that provides a measure of “importance” of words relative to the rest of the network: Harmonic Centrality.

Although our focus is on relations between lexico semantic structure and AoA, it is important to consider any relations in the context of word frequency effects. First, we note there is a relation between AoA and the frequency with which words appear in child directed speech (e.g., Braginsky, Yurovsky, Marchman, & Frank, 2016; Schneider, Yurovsky, & Frank, 2015). However, word frequency in child directed speech may itself be contingent upon the semantic characteristics of words (e.g., Ma et al., 2008). Therefore, any apparent relations between the semantic featural characteristics of words and AoA may result from relations between those characteristics and the frequency with which words appear in child directed speech, placing the locus of the effects in the minds of the speakers or properties of the environment rather than the children’s learning abilities. Second, because our semantic metrics are based on adult generated semantic features, we also consider whether word frequencies in adult directed speech shape the feature norms themselves and indirectly influence relations between semantic structure and AoA. For example, it may be easier to produce certain types of features for words with higher frequency in adult directed speech, obscuring whether any relations between featural characteristics and AoA result from the structure of the feature production task rather than the semantic characteristics of the words. Thus, in study 1 we include exploratory investigations of the role of featural characteristics in determining AoA in the context of word frequency in both child and adult directed speech.

We consider two, non-exclusive, theoretically-informative outcomes. One possibility is that the order of word acquisition may be driven by a single feature type. Such an outcome would be consistent with theoretical proposals that highlight a shift in learning (such as the thematic-to-taxonomic or perceptual-to-conceptual shifts); where early in development certain types of features (functional and perceptual, respectively) dominate, and only later do other feature types begin to play a role. A second possible outcome is that featural characteristics at only certain levels of structure will relate to AoA. For example, it is possible that local- and neighborhood-level featural characteristics will relate to AoA but lexicon-level featural characteristics will not. This would constrain the range of possible cognitive mechanisms through which featural characteristics influence early word learning.

2.1. Methods

2.1.1. Selection of Concepts and Features

2.1.1.1. Early Concepts.: The early-learned concepts were selected from the concrete nouns on the words and sentences form of the MacArthur-Bates Communicative Developmental Inventory (MBCDI:WS; Fenson et al., 1996). The MBCDI:WS includes a parental checklist of common early-produced words. Although this measure of productive vocabulary provides only a limited sample of word knowledge, it has repeatedly associated with a range of group and individual differences in lexical and lexico-semantic processing

(e.g., Borovsky et al., 2016; Fernald et al., 2006; Rämä et al., 2013). Furthermore, while past research had been limited to 130 nouns for which feature norms were available, these analyses use all 359 concrete nouns on the MBCDI:WS.

2.1.1.2. Features.: These analyses use a recently-developed extension (Peters, McRae, & Borovsky, in prep) of the McRae semantic feature production norms (McRae et al., 2005) that now cover the full set of concrete nouns on the MBCDI:WS. We use the subset of 2,417 unique features that appear in the 359 concepts that are in the MBCDI:WS (4,943 concept-feature pairs).

The features are classified according to feature type. Features types are determined as per a taxonomy proposed by Cree and McRae (2003), which categorizes features as functional, perceptual (smell, sound, tactile, taste, visual color, visual form and surface, and visual motion), taxonomic and encyclopedic features. Following prior work (e.g. Hills et al., 2009a; 2009b), the seven types of perceptual features were collapsed into a single feature type in our analyses, resulting in a simplified taxonomy with four feature types (functional, perceptual, taxonomic and encyclopedic).

Finally, it is worth noting a critique of using these features to address the current research question: we are using adult generated semantic features to explore relations in toddlers, who very likely do not have the same semantic representations as adults. This is an important factor which, although tangential to the current project, we explore in supplementary analyses (see supplementary analyses section 3).

2.1.2. Estimating Age of Acquisition—For each of the 359 concrete nouns in the toddler version of the MBCDI:WS, we calculated Age of Acquisition (Wordbank-AoA) using the Wordbank database (Frank et al., 2016). Wordbank is an open repository of MBCDI data, of which we used 5450 administrations of the MBCDI:WS. Using this data, we calculated the proportion of children that knew a certain word for all administrations given to children of a given age ranging from 16 to 30 months. Following Braginsky, Yurovsky, Marchman and Frank (2017), we then calculated the Wordbank-AoA of each word by fitting a logistic curve to the proportions and determining the point at which the curve crossed 0.5.¹ This measure of AoA builds on the standard method for determining AoA from MBCDI proportion data, in which AoA is taken as the first month for which proportions exceed 0.5 (Dale & Fenson, 1996), but also accounts for surrounding data points and patterns across words.

2.1.3. Estimating Word Frequency in Child and Adult Directed Speech—For each of the 359 concrete nouns in the toddler version of the MBCDI:WS, we calculated the natural log of the word's frequency in Child Directed Speech (CDS Frequency) in the CHILDES database (MacWhinney, 2000), using speech directed at children less than or equal to 30 months of age in the North American English collection of chldes-db-version-0.1.0 (Sanchez et al., in prep). The age cut-off was chosen to match the upper age

¹The AoA for MOMMY and DADDY came out negative, so the AoA of the first non-negative word, BALL, was added to their values, thus retaining rank order.

range of MBCDI data used to calculate AoA. As our focus is on conceptual understanding, we collapsed across singular and plural, and across closely related words (e.g., mommy, mama, momma).

For the 348 MBCDI:WS nouns for which data existed, we calculated the natural log of each word's frequency in Adult Directed Speech (ADS Frequency) using the updated version of the SUBTLEX-US corpus with part-of-speech information (Brysbaert & New, 2009; Brysbaert, New, & Keuleers, 2012). As with CDS Frequency, we collapsed across singular and plural, and across closely related words.

2.1.4. Approach to Network Construction—For the neighborhood- and lexicon-level analyses, we constructed graph-theoretic lexico-semantic networks, containing all 359 concrete nouns in the MBCDI:WS, in which nodes represent words and links between nodes represent shared features. To explore the roles of different types of features we constructed separate networks for each of the four feature types. In these networks, following prior precedent (e.g., Engelthaler & Hills, 2017; Hills et al., 2009a; 2009b), the links between concepts were treated as unweighted, undirected edges.

2.1.5. Approaches to Calculating Metrics of Semantic Characteristics

2.1.5.1. Local (Word)-Level Metric.: The local (word)-level analyses focused on the measure *Number of Features* (NoF), which was calculated as the number of semantic features for each of the four feature types. NoF is one way to characterize the semantic richness of a concept (Pexman, Lupker, & Hino, 2002; Pexman et al., 2008). When presented with words with high NoF, adults are faster and more accurate in both lexical decision and semantic categorization tasks than when presented with low NoF words (Grondin, Lupker, & McRae, 2009; Pexman et al., 2002; 2008), and there is tentative evidence that it may be easier to learn words with high NoF (Rabovsky, Sommer, & Abdel Rahman, 2011). Such semantic richness effects emerge in models in the form of stronger attractor representations of word meaning (Plaut & Shallice, 1993) and stronger activation (Rabovsky & McRae, 2014) for words with more features. These models and some behavioral evidence (Grondin et al., 2009) seem to point to a key role in NoF effects for features that are shared across concepts. However, other behavioral and modeling evidence (Cree, McNorgan, & McRae, 2006; Engelthaler & Hills, 2017) indicate that distinctive features, which are not shared with other concepts, are important. Thus, it is worth noting that NoF takes into account distinctive features, and thus may map onto cognitive mechanisms that process such aspects of distinctiveness, while the following higher-level metrics do not.

2.1.5.2. Neighborhood-Level Metrics.: The neighborhood-level analyses focused on two metrics: degree and clustering coefficient. The *degree* (k) of a given word is calculated as its number of edges, or connections to other words. For example, in the first-15-words perceptual network (Figure 2), DOG has a degree of 1 and KITTY has a degree of 5. In the context of lexico-semantic networks, degree is also called number of semantic neighbors or semantic neighborhood density, and has been highlighted in accounts of semantic richness (Pexman et al., 2008), word processing (Mirman & Magnuson, 2008), and word learning

(Storkel, 2009; Storkel & Adlof, 2009). NoF and degree are closely related, but, as noted above, they differ in that NoF provides a fuller representation of the semantic characteristics of a word while degree focuses solely on the aspects of a word's meaning that overlaps with other words. Perhaps due to these differences, NoF and degree account for unique variance in lexical decision response times (Pexman et al., 2008) and may map on to slightly different cognitive mechanisms. For example, degree is used to characterize the relative importance of nodes in a network regarding their potential communication activity, which, in the context of lexico-semantic networks, nicely corresponds to the mechanism of spreading activation.

The *clustering coefficient* (CC) of a given word quantifies the degree to which the word's neighborhood is interconnected, and is defined as the ratio of the number of links between the neighbors of a word and the maximum possible number of links between the neighbors (Watts & Strogatz, 1998). Clustering coefficient ranges from 0, when none of a word's neighbors are connected, to 1, when all a word's neighbors are connected.² For example, in the first-15-words perceptual network, the clustering coefficient of both SHOES and CAR is 1, because all their neighbors are interconnected, while the clustering coefficient of KITTY is .4, because only some of its neighbors are connected. As with degree, clustering coefficient may relate to patterns of spreading activation in a lexico-semantic network. For example, Mirman and Magnuson (2008) showed that processing is inhibited for words with dense near neighborhoods but facilitated for words with dense far neighborhoods. They defined near neighbors as words that share at least half of their features, and far neighbors as those that share fewer features. As a direct result of sharing such a high proportion of features, near neighbors are more likely to share neighbors, and consequently have a higher clustering coefficient, than far neighbors. Vitevitch, Ercal, and Adagarla (2011) demonstrated a similar result using network simulations. Thus, clustering coefficient may relate to patterns of inhibition and facilitation in word and concept processing that depend on spreading activation feedback between neighbors.

2.1.5.3. Lexicon-Level Metric.: The lexicon-level analyses focus on a measure of “importance” of words relative to the rest of the network, harmonic centrality. Harmonic centrality is a modification of closeness centrality (Freeman, 1978) that resolves problems resulting from infinity values in the denominator (Rochat, 2009). Closeness centrality measures in network graph theory are designed to rank words in importance relative to their closeness to, and thus ability to communicate efficiently with other words in the network. Thus, in the context of lexico-semantic networks, harmonic centrality may relate to the efficiency with which activation spreads from a node across the full range of the network.

The *harmonic centrality* (HC) of a given word (i) is defined as the sum of the reciprocals of the pairwise distances between that word and all other words in a network, divided by the number of words in the network (n) minus 1 (Rochat, 2009):

²Common practice in the field is to replace undefined values with either zeros or ones (e.g., Costa, et al., 2007). Results for the simple neighborhood level model remained relatively unchanged for each method in the base models, but were different for models with word frequency included. One option is to simply remove concepts with undefined values, but this results in losing 271 data points (with 101 undefined values for encyclopedic, 126 for functional, 35 for perceptual and 162 for taxonomic). Thus, to avoid strong effects on coefficients due to high leverage artificial values at either tail of the CC distributions while maintaining the full range of data points, we replaced undefined values with the means of their respective distributions.

$$HC_i = \frac{\sum_{j=1}^n \frac{1}{distance_{ij}}}{n-1}$$

Where the *distance*, or shortest path length, between two words (*i* and *j*) is defined as the smallest number of links that must be traversed to get from one node to the other. For example, in the first-15-words perceptual network the distance between DOG and KITTY is 1, while the distance between DOG and DUCK is 3. In cases where there is no connecting path, such as between DOG and DADDY, the distance is equal to infinity. Traditional closeness centrality, where the sum is in the denominator, cannot be used in such networks because an infinity value in the sum results in closeness values of 0 for all nodes. Harmonic centrality ranges from 0, for a node with no links, to 1, for a node that is directly linked to all other nodes in the network.

As is apparent in Figure 2, degree and harmonic centrality are similar, but not identical, because harmonic centrality includes information regarding words that are outside its immediate neighborhood. For example, NOSE and BABY both have a degree of 4, but NOSE has a higher harmonic centrality (.534) than BABY (.512) because its connection to highly connected BALLOON results in a lower average distance (and higher inverse average distance) to the other words in the network.

2.2. Study 1 Results

As described above, we sought to explore whether and how featural knowledge of early acquired nouns and their corresponding lexico-semantic networks relates to lexical Age of Acquisition (AoA). We explore this question by measuring the interaction of featural connections at three levels of lexico-semantic structure: local (word)-, neighborhood-, and lexicon-levels. In our preliminary inspection of the data, we noted complex interactions between levels of lexico-semantic structure and feature type (see supplementary analyses section 1). Therefore, we sought to clarify the patterns in the data by carrying out separate analyses at each level of lexico-semantic structure in three separate multivariate models, to address questions regarding which feature types and levels of network structure best predict AoA in early lexical acquisition. Furthermore, we conducted exploratory investigations of the role of featural characteristics in determining AoA in the context of word frequency in both Child Directed Speech (CDS) and Adult Directed Speech (ADS). Thus, at each level of analysis, we conduct four multivariable linear regression models, using scaled and centered variables (so that coefficients can be compared across metrics and models), in which Wordbank-AoA is a function of (1: Base Model) that level's set of featural characteristics, (2: +CDS Model) featural characteristics and word frequency in child directed speech, (3: +ADS Model) featural characteristics and word frequency in adult directed speech, and (4: Full Model) featural characteristics and word frequency in both child and adult directed speech.

2.2.1. Local (Word)-Level Analyses—We first modeled Wordbank-AoA as a function of the local (word)-level featural characteristic: Number of Features (NoF), for each of the four feature types. The Base Model significantly associated with AoA, $F(4,354) = 7.60, p < .$

001, $R^2_{\text{adj}} = .07$, with perceptual and taxonomic NoF both contributing significantly to the model (coefficient values for all local-level analyses are presented in Table 1). This outcome suggests that words that have more perceptual features are learned earlier by toddlers.³ Secondly, the outcome suggests that word-level taxonomic features are also related to AoA, though less robustly than perceptual features.

Next, we explored the role of local (word)-level featural characteristics in the context of word frequency in child and adult directed speech. The +CDS Model was significant, $F(5,353) = 57.45$, $p < .001$, $R^2_{\text{adj}} = .44$, with perceptual NoF and CDS frequency both significantly predicting AoA. The results of the +ADS Model indicated four predictors – functional, perceptual, and taxonomic NoF and ADS frequency – significantly predicted AoA, $F(5,342) = 16.12$, $p < .001$, $R^2_{\text{adj}} = .18$. Finally, the Full Model also significantly associated with AoA, $F(5,341) = 59$, $p < .001$, $R^2_{\text{adj}} = .50$, with perceptual NoF, CDS frequency, and ADS frequency all contributing significantly to the model. Across these three models, although the size of the perceptual NoF effect decreases upon the addition of CDS frequency, the relation between perceptual NoF and AoA remains significant even with *both* CDS and ADS frequencies added to the model. These results suggest that perceptual NoF accounts for variance in the age at which words are learned above and beyond what is accounted for by the frequency with which words appear in either child or adult directed speech. However, the fact that taxonomic NoF loses significance in the +CDS and Full models indicates it may be accounting for variance in the Base and +ADS models that is better explained by word frequency in child directed speech.

2.2.2. Neighborhood-Level Analyses—We next modeled Wordbank-AoA as a function of the neighborhood-level featural characteristics: Degree and Clustering Coefficient (CC), for each of the four feature types. The base neighborhood-level model significantly associated with AoA, $F(8,350) = 2.21$, $p = .03$, $R^2_{\text{adj}} = .026$, however, only Degree-Perceptual contributed significant variance to this model (coefficient values for all neighborhood-level analyses are presented in Table 2). This analysis indicates that words with more perceptual links to other words are learned earlier than words with lower perceptual degree.

For the exploration of the neighborhood-level featural characteristics in the context of word frequency in child and adult directed speech, we first found the +CDS Model was collectively significant, $F(9,349) = 32.79$, $p < .001$, $R^2_{\text{adj}} = .44$, with perceptual and functional degree and CDS frequency significantly predicting AoA. The +ADS Model was also significant, $F(9,338) = 7.94$, $p < .001$, $R^2_{\text{adj}} = .15$, with two significant predictors of AoA: perceptual degree and ADS frequency. Finally, in the Full Model, encyclopedic degree and both ADS and CDS frequency contributed to a significant association with AoA, $F(10,337) = 35.96$, $p < .001$, $R^2_{\text{adj}} = .50$. Unlike the local (word)-level models, the relation between perceptual degree and AoA does not remain significant when *both* CDS and ADS frequencies are added in the full model. These results suggest that the relation between perceptual degree and AoA is not as robust as for perceptual NoF.

³Further exploration indicates the relation between perceptual NoF and AoA in the Base Model is largely driven by visual-form and surface features, with a smaller role for visual-motion features. For details see supplementary analyses section 2.

2.2.3. Lexicon-level Analyses—Finally, we modeled Wordbank-AoA as a function of the lexicon-level metric Harmonic Centrality (HC), for each of the four feature types. The results of the regression indicated that the Base Model was collectively significant, $F(4,354) = 3.42, p = .009, R^2_{adj} = .026$, with perceptual and taxonomic harmonic centrality both significantly predicting AoA (coefficient values for all lexicon-level analyses are presented in Table 3). This analysis suggests that earlier-learned words contain perceptual and taxonomic connections that make them closer to other words in the lexicon.

In our exploratory analyses looking at the lexicon-level featural characteristics in the context of word frequency in child and adult directed speech we first found the +CDS Model significantly associated with AoA, $F(5,353) = 57.68, p < .001, R^2_{adj} = .44$, with perceptual and functional harmonic centrality and CDS frequency contributing significantly to the model. Next, in the +ADS Model, perceptual HC and ADS frequency contributed to a significant association with AoA, $F(5,342) = 14.98, p < .001, R^2_{adj} = .17$. Finally, the Full Model was also significant, $F(6,341) = 58.72, p < .001, R^2_{adj} = .50$, but with only two significant predictors of AoA: CDS and ADS frequency. Like the neighborhood analyses, the relation between perceptual HC and AoA remains significant across both the +CDS and +ADS models but not in the full the model.

2.3. Study 1 Discussion

Our analyses sought to explore how the semantic characteristics of word meanings contribute to the order of early-word learning. We explore this question by measuring the interaction of two factors: (1) feature type: encyclopedic, functional, perceptual or taxonomic, and (2) whether feature type interacted with different levels of lexico-semantic structure. Furthermore, we conducted exploratory analyses investigating these interactions in the context of two potentially confounding variables: word frequency in child directed speech and word frequency in adult directed speech.

Considering first the question of structure, the base models of local (word)-level structure explain the most variance in AoA, compared to more complex measures that consider lexical connectivity to immediate and more distant neighbors, which were relatively poorer predictors of lexical acquisition (though still significantly associated). Nonetheless, it is notable that despite these differences across structure, perceptual features contribute significant variance in the base linear models of all levels of structure. Furthermore, at the local (word)-level, perceptual Number of Features contributed significantly even after the addition of word frequency metrics for both child and adult directed speech. This contrasts with taxonomic featural characteristics, which no longer contribute significantly in models with word frequency in child directed speech. The consistent perceptual feature pattern in combination with decreasing model fits across levels of structure is consistent with the possibility that the relations between AoA and higher-level structure might only arise due to their relations with the word-level feature characteristics. This explanation is further supported by the results of “mega” model analyses, containing all metrics from study 1 (supplementary analyses sections 4 and 5), in which perceptual Number of Features was a significant predictor of AoA, but perceptual degree and harmonic centrality were not. Alternatively, this pattern of results could also be due to non-linear relations between AoA,

vocabulary size, and featural characteristics at higher levels of structure – a possibility we explore in study 2.

However, the analytical approach in study 1 was not able to explore these potential nonlinear relations between AoA and lexical structure as a function of vocabulary growth. This limitation arises because we calculated network metrics using the full lexicon networks rather than those that would exist at the point in development when a word is acquired. In other words, we might expect different patterns in lexical connectivity when children only know several words, as compared to when they know many words. However, the methods employed in study are not sensitive to changes in lexical size over time, and thus cannot distinguish between these possibilities. We approach the research question from a new direction in study 2, using a novel method that allows us to avoid this limitation.

3. Study 2: Analyses of longitudinal models of normative lexico-semantic network growth using local-, neighborhood-, and lexicon-level measures of network structure

Study 2 differs from study 1 in two main ways: (1) We explicitly model developmental changes in lexico-semantic structure as a function of the lexicon at the points in development when each word is learned and (2) We measure changes in the overall lexical network structure over time, rather than looking at relations with metrics centered on individual words – in other words, in study 1 the unit of analysis is the word, while in study 2 it is the entire network. This approach allows us to measure how semantic features influence lexico-semantic structure as a network grows over time, which in turn allows us to examine whether the patterns of network structure that related to AoA in study 1 linearly hold true across the range of development or if there are non-linear patterns that were obscured. Building on prior research (e.g., Hills et al., 2009b), we create normative growth networks in which words are added to the network following their Wordbank-AoA rank order. Then, we compare the trajectories of featural characteristics across development to networks that grow in a random order.

Our approach builds on the standard practice in network science of determining if networks have large-scale structure by comparing normative lexical growth networks, where words are added per their Wordbank-derived AoA, to random lexical growth networks, where new words are added in random orders. This comparison allows us to explore longitudinal differences that result precisely due to the order in which words are learned, building on the analyses from study 1. To determine time windows in these developmental trajectories where different features influence network structure, we use a non-parametric cluster analysis technique commonly used in EEG (e.g. Groppe, Urbach & Kutas, 2011a; Maris & Oostenveld, 2007) and eye-tracking (e.g. Barr, Jackson, & Phillips, 2014; Von Holzen & Mani, 2012). This technique identifies clusters of significant difference between the normative and random trajectories, which in turn allows us to explore the timing of the roles of different types of features in early lexico-semantic development.

Following study 1, our analyses in study 2 vary along two dimensions – (1) feature type and (2) level of structure within a network (i.e. local (link)-, neighborhood-, and lexicon-level). First, the analyses of feature-types parallel our approach in study 1, which vary by feature type (i.e., encyclopedic, functional, perceptual, and taxonomic). Second, the levels of structure also parallel those of study 1, but with an important difference: in study 1 the metrics focused on individual words and the levels referred to the different grain sizes of structure that described how those words connected to the network, but in study 2 the focus is on the structure itself and the levels of structure refer to the grain size of the information used to calculate metrics. At the local (link)-level, we focus on a summary of the featural substructure of individual links defined by link count. At the neighborhood-level, we focus on density and clustering, which are apparent at the neighborhood-level. At the lexicon-level, we focus on two descriptions of large-scale structure: mean distance and a quantitative metric of small-world-ness.

Given the analytic plan to explore the impact of features, structure and developmental change in lexical growth networks, we consider a number of non-exclusive, theoretically interesting outcomes where normative lexical network growth may show systematic differences from random growth networks – indicating relations between the featural characteristics and the normative order in which words are learned. For example, normative vs. random growth may differentiate according to (1) specific features at all points in development, (2) at specific levels of structure at all points in development, and (3) at specific features and/or structure at some, but not all, periods of development. We consider several likely outcomes within each of these patterns that might lend further support to the patterns observed in study 1, where perceptual features, and taxonomic to a lesser degree correlated most consistently with lexical AoA. For the first potential pattern (systematic growth driven by specific features across development), we would expect that normative perceptual feature networks should show enhanced structure compared to random perceptual networks. This pattern would bolster theoretical accounts that emphasize the importance of perceptual features in early lexico-semantic development, and provide evidence that the locus of the effect lies in the abilities of the learner – perhaps as an attentional bias that prioritizes particular kinds of perceptual features (e.g., the shape bias; Landau et al., 1988) – rather than in the environment. The second outcome (systematic growth only at specific levels of structure), might result in systematic perceptual growth at the local (link)-level, where we saw the strongest effects in study 1. This potential result might suggest that perceptual properties relate to lexical growth, in that individual perceptual properties may simply be easier or more salient in word learning, rather than driving large-scale lexical-semantic structure, per se. Regarding the third potential outcome (developmental variation in features and structure), we might expect to find systematic relations between perceptual features and lexical growth at the earliest points of development, while other more weakly associated features, such as taxonomic properties, relate to growth later in development. This account might be most consistent with accounts that posit a perceptual-to-conceptual shift in early lexical development (e.g., Quinn & Eimas, 1997). Alternatively, we might also see a developmental shift in structure, such that lower-level local structure might be emphasized earlier, while neighborhood and lexicon-level patterns show systematic growth at later points

in network growth, mirroring the better model fit parameters for the base local (word)-level model in study 1.

3.1. Methods

3.1.1. Selection of Concepts and Features, and Estimation of AoA—In study 2, we use the same early concepts and features as in study 1. We also use the same method for estimating AoA as in study 1.

3.1.2. Approaches to Network Construction—For the local (link)-level analyses, we construct a sequence of 359 graph-theoretic lexico-semantic networks in which nodes represent words and links between words represent shared features. Crucially, in these networks, the links between words are *not* collapsed. Thus, if two words share four features, there will be four links connecting them.

For the neighborhood- and lexicon-level analyses, we construct four sets of sequences of 359 graph-theoretic lexico-semantic networks like those used in the neighborhood- and lexicon-level analyses in study 1, in which the links between words *are* collapsed. To explore the roles of different types of features we constructed separate sets of networks for each of the four feature types.

3.1.3. Approaches to Calculating Metrics

3.1.3.1. Local (link)-level Metric. For the local (link)-level analyses, we focus on a metric that summarizes link-level information: *link count*, defined as the number of links of a given feature type that are in the network. This metric is corollary to the Number of Features metric used in study 1, in that it is basically the number of features that are shared by pairs of words. This measure can be conceived as characterizing the amount and type of overlap in semantic representations across the network, which in turn may relate to cognitive mechanisms such as the potential for spreading activation of different types of meaning among known words or the likelihood that a certain type of feature will be recognized in a novel object. One shortcoming of link count is that it does not account for the total number of features of a given type that could possibly result in links for any given pair of concepts. This is important since there are significant differences in the distributions of feature types both across concepts (Cree & McRae, 2003) and more generally, with perceptual features outnumbering all the other types of features (Devereux et al., 2014; McRae et al., 2005; Peters et al., in prep). Therefore, we also use a measure of *mean normalized link weight*, defined as the mean of weighted links of a given feature type that have been normalized for the total number of unique features, of the given feature type, across the two concepts they connect. For example, BIRD and KITTY share 1 perceptual feature, and thus 1 perceptual link, and have a total of 16 unique perceptual features (8 for BIRD and 9 for KITTY, minus the 1 they share), so for this pair the weighted link is calculated to be $1/16 = .0625$. The normalized link count for the network is the mean of all such weighted link values in the network, which is equivalent to the percentage of semantic representations that are shared between words out of what could possibly be shared across the whole network.

3.1.3.2. Neighborhood-level Metrics.: For the neighborhood-level analyses, we focus on two metrics: density and global clustering coefficient. The *density* of a network is defined as the number of links divided by the maximum number of links possible for a network with the same number of words, and is equivalent to the mean degree of a network, normalized for the number of nodes. This results in a number that ranges from 0 to 1, which is equivalent to the probability that any two words will be connected. The *global clustering coefficient* (*GCC*), as with the clustering coefficient used in study 1, characterizes the amount of clustering in a network. However, rather than being defined in terms of the neighborhoods of individual words, global clustering is defined as the total number of closed triangles (e.g. KITTY-BABY-BIRD in Figure 2) in a network divided by the number of connected triples (Newmann, Moore, & Watts, 2000). As with the neighborhood-level metrics in study 1, density and global clustering coefficient may relate to the cognitive mechanism of spreading activation and, consequently, patterns of inhibition and facilitation in word and concept processing that depend on feedback between neighbors (Mirman & Magnuson, 2008; Vitevitch et al., 2011).

3.1.3.3. Lexicon-level Metrics.: For the lexicon-level analyses, we focus on two metrics: mean distance and the metric of small-world-ness. The *mean distance* ($\langle d \rangle$) of a network is defined as the mean of the distances, or shortest path-lengths, of all pairs of words in a network for which a connecting path exists. Mean distance and global clustering coefficient have both been of interest to researchers exploring small-world properties in semantic networks (e.g. Beckage et al., 2011; Hills et al., 2009a; Steyvers & Tenenbaum, 2005). A network is qualitatively defined as having *small-world properties* if it has *both* dense neighborhood-level clustering and lexicon-level structure that results in short distances between words (i.e. high global clustering coefficient and low mean distance; Watts & Strogatz, 1998). In the context of lexico-semantic networks, such structure has been argued to have the functional purpose of increasing the efficient spread of activation across the full range of the network, which in turn results in fluent processing and word retrieval (Griffiths, Steyvers, & Firl, 2007; Steyvers & Tenenbaum, 2005). A quantitative *metric of small-world-ness* (*SW*; Humphries & Gurney, 2008) that builds on the qualitative definition of Watts and Strogatz (1998) defines small-worldness as:

$$SW = (GCC/GCC_{ER}) / (\langle d \rangle / \langle d \rangle_{ER})$$

Where GCC_{ER} and $\langle d \rangle_{ER}$ are the average global clustering coefficient and average mean distance of 2,000 equivalent (i.e. same number of words and links) Erdős-Rényi random graphs (Erdős & Rényi, 1959) for the network being tested. Small world networks are then defined as those for which *SW* is significantly greater than SW_{ER} , the value of *SW* for an equivalent ER-graph, which generally holds true for $SW > 1$.

3.1.4. Modeling Lexico-Semantic Network Development

3.1.4.1. Modeling Normative Network Development.: To model normative lexico-semantic network development we create sets of feature networks where words are incrementally added per their AoA rank. This growth procedure is repeated across each of

the five networks discussed in section 3.1.2, resulting in five sets of progressive normative growth networks with 359 steps.

3.1.4.2. Modeling Random-AoA Network Development.: To model random-AoA lexico-semantic network development we create sets of feature networks where the same 359 early-learned words used to model normative development are incrementally added per a randomized AoA ranking. This growth procedure is repeated 2000 times (as recommended by Groppe, Urbach, & Kutas, 2011b for the cluster analysis technique, see below) for each of the five networks discussed in section 3.1.2. The mean metric values for each step, across iterations, are taken to represent patterns in lexico-semantic network development that occurs without the influence of Wordbank-AoA. It is important to note that by using the same 359 words across normative and random-AoA models we are inherently constraining the networks to become equivalent at the end of development.

3.1.5. Using Cluster Analyses to Compare Trajectories—We compared normative vs random-AoA development using a non-parametric cluster analysis technique (e.g. Groppe, Urbach, & Kutas, 2011a; Maris & Oostenveld, 2007). This statistical technique is ideal for determining the point-wise differences between two time-series while correcting for multiple comparisons. “Cluster t-statistics” are calculated by summing across clusters of sequential point-wise t-values that exceed a pre-specified threshold. For our comparisons, we grouped measures into bins of 10 sequentially ordered words (e.g. the first ten words according to Wordbank-AoA were in Bin 1 in the normative trajectory), and defined the threshold as being significantly different as determined by a two-tailed t-test with significance level of .05. This procedure was then repeated for all 2,000 iterations, to create a rank-ordered cluster t-statistic distribution of the largest clusters from each of the 2,000 trajectories from the random comparison network. Finally, any cluster t-statistics from the comparisons of normative vs random-AoA were compared against the distribution, resulting in non-parametric Monte Carlo p-values (equivalent to the percentage of the number of random clusters greater than the cluster t-statistic of interest).

3.2. Study 2 Results and Discussion

Our analyses address the question: Do the trajectories of featural characteristics of lexico-semantic networks differ from random over the course of normative development? These featural characteristics differ across two dimensions – (1) feature types and (2) level of analysis (i.e., local (link)-, neighborhood-, and lexicon-level). Contrary to study 1, our use of the non-parametric cluster analysis method necessitates a separate analysis for each comparison, but analyses will be grouped by level of analysis. For each analysis, we first qualitatively characterize the shape of the normative and random-AoA trajectories, then present statistical comparisons with random development, and finally, we briefly discuss the findings.

3.2.1. Local (link)-Level Analyses

3.2.1.1. Link Count Trajectories.: We first illustrate the normative trajectories across all four feature types in figure 3 for (A) link count and (B) mean normalized link weight. Figure 3A qualitatively indicates that perceptual features dominate links between concepts in early

development. Figure 3B suggests that advantage for perceptual features remains, even when controlling for the total number of available features. Interestingly, while taxonomic features do not stand out from encyclopedic and functional features in the graph of raw count (Figure 3a), taxonomic links between concepts do appear with greater frequency than either functional or encyclopedic features once normalized for their relatively small numbers (Figure 3b).

In Figure 4 we present the normative and random-AoA link count trajectories for each of the four feature types, in bins of 10 words. Both normative and random-AoA feature count trajectories start at 0 and converge at the final time step because the full network of 359 words is equivalent for both trajectories. We compared the trajectories using the cluster analysis method described in section 3.1.3. There were no significant clusters of sequential significant differences for either encyclopedic or functional features, indicating that link counts were not driven by systematic differences between normative and random growth for these two feature types. On the other hand, cluster analyses did reveal that link counts for normative growth networks built using both perceptual and taxonomic features exceeded random growth over nearly the entire period of development, encompassing bins 2 through 36 for both feature types (perceptual: *cluster-T* = 346.05, *monte-carlo* p = .001, taxonomic: *cluster-T* = 538.99, *monte-carlo* p = .016). These results naturally follow from the relations between perceptual and taxonomic Number of Features and AoA, in study 1, given that (1) having more features equates to having more shared features with other concepts and (2) the linear relations hold true for dynamic growing networks.

3.2.2. Neighborhood-Level Analyses

3.2.2.1. Density Trajectories.: The normative and random-AoA density trajectories for each of the four feature types are illustrated in figure 5. Qualitatively, the density trajectories have relatively flat slopes, meaning density is largely unchanging across development, but perceptual networks are denser (~.15) than the rest (~.025). Quantitatively, density trajectories for networks constructed from encyclopedic and functional features did not significantly differ from random. However, the density trajectories for lexical growth networks with perceptual and taxonomic features did exceed that of random lexical networks across nearly the entire developmental trajectory, ranging from bin 2 through 36 (perceptual: *cluster-T* = 1712.12, *monte-carlo* p = .002, taxonomic: *cluster-T* = 1273.82 *monte-carlo* p = .04). These results indicate that over the course of normative development, both perceptual and taxonomic networks consistently have more links than would be expected to happen randomly – as should be expected given the evidence in study 1 that words with more perceptual and taxonomic links are typically learned earlier.

3.2.2.2. Global Clustering Coefficient Trajectories.: Figure 6 shows the normative and random-AoA global clustering coefficient trajectories for each of the four feature types. Qualitatively, as with density, the global clustering coefficient trajectories have relatively flat slopes, again meaning the degree of clustering is largely unchanging across development. However, functional and taxonomic networks demonstrate more clustering (~.85-.9) than encyclopedic and perceptual networks (~.5-.6). Quantitatively, functional, perceptual and taxonomic feature networks did not show systematic differences in clustering from random

growth networks along any point in lexical development. However, lexical growth networks with encyclopedic features tentatively demonstrated less clustering than random, ranging from bin 8 through 29 ($cluster-T = 1334.46$, $monte-carlo p = .057$). In other words, over the course of development, normative networks show similar patterns of clustering as networks with randomly chosen words. Furthermore, the lack of significant nonlinear patterns corroborates the with the null relations between clustering coefficient and AoA, in study 1.

3.2.3. Lexicon-level Analyses

3.2.3.1. Mean Distance Trajectories.: In Figure 7 we present the normative and random-AoA mean distance trajectories for each of the four feature types. Qualitatively, both normative and random-AoA trajectories generally start off low and quickly rise to a peak in the first half of the trajectory, before steadily decreasing for the remainder of the trajectory. The low starting point is due to the fact that mean distance ignores disconnected nodes, which means that early in development, when there are relatively few links, pairs or small clusters of nodes that are directly connected dominate. Then, as these small groups begin to link up, forming larger clusters, mean distance increases. Finally, as more complex structure begins to develop the trend reverses and then mean distance stabilizes. While the peaks of the perceptual trajectories occur sooner and are less than for the other feature types, the only trajectories not following this pattern are the taxonomic trajectories, which slowly but steadily increase across the full range of development. Quantitatively, encyclopedic, functional and taxonomic feature networks did not differ significantly in mean distance from random growth networks along any point in lexical development. However, the mean distance for lexical growth networks with perceptual features was significantly less than for random, ranging from bin 3 through 31 ($cluster-T = -1769.9$, $monte-carlo p = .001$). In other words, for a portion of development, the number of perceptual links separating words is significantly less (i.e., the words are closer) than would be expected to happen randomly.

3.2.3.2. Metric of Small-World-ness Trajectories.: Finally, the normative and random-AoA metric of small-world-ness (SW) trajectories for each of the four feature types are illustrated in figure 8. Qualitatively, the trajectories for all feature types are all numerically greater than 1 through nearly the entire length of the modeled trajectories, indicating that, irrespective of the order in which words are acquired, all feature types develop small world properties from an early point in lexical development. Quantitatively, functional, perceptual and taxonomic feature networks did not show systematic differences in the degree of small-world structure from random networks along any point in lexical development. However, lexical networks with encyclopedic features marginally demonstrated less small-world structure than random ranging from bin 6 through 34 ($cluster-T = -542.03$, $monte-carlo p = .079$).

4. General Discussion

Theoretical accounts of early lexico-semantic development have debated that different kinds of word meaning might take priority in early learning and development, alternatively highlighting the initial importance of perceptual and functional semantic features. Using a large database of toddlers' normative word growth in combination with a recently developed

set of semantic feature norms for early-acquired nouns, we sought to contribute to these theoretical accounts of early lexical development by determining whether different semantic properties of word meanings associate with the order in which words are normatively learned from 16- to 30-months. Our results are clear: perceptual features were repeatedly identified as relating to the order of early lexical development across a range of measures of lexical structure, and taxonomic features, to a lesser degree, also related to the pattern of early lexico-semantic growth.

The two studies provide converging evidence via two approaches. First, in study 1, taking the word as the unit of analysis, we explored the linear relations between the AoA of words and words' local (word)-, neighborhood-, and lexicon-level featural characteristics. Furthermore, exploratory analyses investigated these relations in the context of word frequency in both child- and adult-directed speech. Then, in study 2, taking the entire lexicon as the unit of analysis, we modeled noun-feature growth networks to explore nonlinear relations between the normative order in which words are learned and the developmental trajectories of local (link)-, neighborhood-, and lexicon-level featural characteristics over the course of early word learning. These two approaches have contrasting strengths and weaknesses and different units of analysis; thus the results have different implications for interpretation regarding the questions we explored, namely: (1) Whether specific semantic features (e.g., perceptual or functional) contribute to early lexico-semantic development, (2) Whether different levels of structure associate with early growth, and (3) Whether features and structural properties of early lexico-semantic growth vary and interact across time.

With respect to the first question, regarding relations between specific semantic properties and lexical growth, the results indicate that perceptual features have an important role in early lexico-semantic development. In study 1, perceptual featural characteristics of words were the most robust predictors of AoA for all levels of structure, and in study 2, we found evidence of systematic (non-random) growth in perceptual feature networks for all levels of structure. Taxonomic featural characteristics also predicted AoA in study 1 and taxonomic feature networks demonstrated systematic growth at the local and neighborhood-levels in study 2. Although taxonomic associations with AoA were weaker in comparison to perceptual features, this finding is particularly interesting when considering that taxonomic features have been excluded from past work using feature-based lexico-semantic networks (e.g., Hills et al. 2009a, 2009b; Stella, Beckage, & Brede, 2017; Yurovsky, Bion, Smith, & Fernald, 2012). The decision to exclude taxonomic features has been justified based on arguments that such features are unlikely to be part of children's conceptual representations. However, our results indicate the taxonomic featural characteristics of words and networks are related to learning from nearly the very beginning of early word learning, suggesting that early lexico-representations may encode for hierarchical relations between word meanings. This "early taxonomic encoding" explanation treats taxonomic features equivalently to perceptual features; however, many researchers consider taxonomic features to be "different from information regarding other types of features" (McRae et al., 2005, pg. 552). Indeed, taxonomic features convey relational properties that describe higher order structural characteristics of words within a network (e.g., Lucariello, Kyratzis, & Nelson, 1992). This relational information might arise from correlated perceptual features among words (Hills et

al., 2009b), akin to Rosch and Mervis' (1975) conceptualization of family resemblances, suggesting that taxonomic relations among words may be an emergent property of networks with significant perceptual feature overlap. A third explanation for the relation between taxonomic features and AoA is that, rather than taxonomic features influencing word learning directly, they instead indirectly influence it via a relation with word frequency in child directed speech. In other words, rather than the locus of the taxonomic effect taking place in the minds of toddlers as they learn words, it may instead be in the minds of the adults creating the language learning environment (Hills et al., 2009a; Hills et al., 2010). Indeed, this explanation is supported by the results of our exploratory analyses in study 1 of AoA as a function of local (word)-level metrics and word frequency in child directed speech, in which, although perceptual Number of Features did still significantly contribute to the full model, taxonomic Number of Features was no longer significant. Finally, it should also be noted that across a variety of lexico-semantic modeling approaches, functional and encyclopedic features did not associate with patterns of early lexical acquisition. When contrasting this pattern with the robust perceptual feature findings, our results are most consistent with perceptual-to-conceptual accounts of early lexico-semantic development (Quinn & Eimas, 1997; Rakison & Lupyan, 2008; Sheya & Smith, 2006; Smith et al., 1996; etc.).

Our second question concerned whether and how different levels of structure (at local-, neighborhood- and lexicon-levels) predicted AoA. At the outset of the study, we posited three potential relations between structure and AoA. The first was that words with more perceptual features would be learned earlier, and that this relation would extend to higher levels of structure. The second pattern we posited was that words with more perceptual features would be learned earlier, but that this relation would NOT extend to higher levels of structure. Finally, the third was that the number of perceptual features would NOT relate to the order in which words are learned, but higher-level perceptual structure would. Furthermore, we noted that in this final case we might also expect relations between AoA and other kinds of features, namely functional and taxonomic features. In study 1, local (word)-level perceptual characteristics robustly related to AoA (even when accounting for word frequency in both child and adult directed speech), while neighborhood- and lexicon-level characteristics were only weakly related to AoA. We consider two explanations for this latter effect. The first possibility is that lower-level structure could propagate to higher level structure in early lexico-semantic networks. This result lies somewhere between the first two potential outcomes reiterated above; in that higher-level structure across the lexicon is built by word-level networks, but the relation is imperfect, such that features of early-learned words sometimes, but not always, overlap with other concepts. This pattern suggests that young children might be best prepared to learn words with many perceptual features, but that learning may not necessarily be supported by selecting items in perceptually dense semantic neighborhoods. Indeed, as discussed in supplemental analyses section 5, semantic distinctiveness (when words have less overlap with other word meanings) can also associate with AoA. Alternatively, potential non-linearities between higher-level structure and AoA that are not easily captured by linear regression modeling approaches might drive these differences between higher- and lower-level structure. Fortunately, the non-parametric growth modeling approach in study 2 makes it possible to identify non-linear relations. In

study 2, we found evidence for systematic (non-random) growth in perceptual feature networks for all levels of structure, across nearly the full range of development. In concrete terms, toddlers do not simply learn words at random: they prioritize developing perceptual linkages among word meanings. The fact that the patterns of systematicity are nearly identical across levels of analysis lends support to the first explanation above. If, instead, relations between AoA and local-level lexico-semantic characteristics were not propagated to higher levels, we might expect either no systematic differences at higher-levels or periods of systematic differences that differ from lower-levels. Thus, the pattern of results suggests that large-scale structure does not play a driving role in early lexico-semantic development, but rather arises out of the ease of individual word learning, which is in turn related to the richness of the perceptual characteristics of individual words (i.e., Number of Features). Across the two studies, the pattern of results falls somewhere between the first two examples, and thus provides tentative support for both attentional accounts of early word learning that point to the importance of perceptual correlations among objects to drive word learning (e.g. Smith et al. 2002) and those that highlight the importance of distinctiveness (Engelthaler & Hills, 2017). Distinguishing between these possibilities necessitates future empirical investigations.

Lastly, we were interested in how the feature types and level of lexico-semantic structure developed across time. In study 2, for perceptual and taxonomic growth networks, differences between normative and random trajectories extended for nearly the full length of the developmental period, and were in directions consistent with what would be expected if the patterns of relations in study 1, between perceptual and taxonomic featural characteristics and AoA, hold true across the range of development in the context of dynamic growth networks. In other words, these results do not provide evidence for a discrepancy between patterns of featural characteristics in the full lexicon network and networks at the points in time when words are learned. Thus, these results allow for the possibility that the perceptual and taxonomic feature effects are driven by the lexico-semantic networks of the learners, and that these features are encoded and influence word representations from the very onset of lexical development.

These findings have important applied and theoretical implications for early lexical development. Broadly, our findings suggest that perceptual features and linkages among word concepts may support early word learning. For example, the local-level findings in both experiments highlight that concepts with more perceptual features are acquired (on average) earlier across normative development. These findings suggest concepts that have more perceptual features should be easier to learn relative to concepts that have fewer perceptual features. Such a finding may be generally consistent with findings that concepts with greater concreteness and iconicity are acquired earlier (e.g., Gilhooly & Logie, 1980; Perry et al., 2015), and are also broadly consistent with work indicating that objects with greater perceptual salience draw the attention of, and subsequently have labels mapped to them by, young learners (Pruden et al., 2006). However, while this previous work was with 10-month-olds, and the effect appeared to diminish quickly with age, our focus here has been on 18 to 30-month-olds. Furthermore, the effect of perceptual features on word learnability – and the resulting early abundance of perceptual features in the semantic network – persists well past the point that such effects, indicating perceptual features may

continue to play an important role in semantic processing well past any early learnability effect. However, more empirical research is needed to directly contrast how different types of features may specifically support lexical acquisition in young children and whether there are any subsequent longer-term effects on processing. Intriguingly, our findings have additional implications for early word learning that extend beyond the featural composition of individual concepts. Neighborhood-level effects of perceptual degree, in study 1, and perceptual density, in study 2, indicate that concepts that have more perceptual features that are shared by multiple concepts – and thus have more perceptual links – should be easier to learn than concepts with fewer shared features. Interestingly, research by Grondin and colleagues (2009) shows a similar pattern of effects; they showed that adult conceptual processing is facilitated for concepts with more perceptual features, and in particular more shared perceptual features. Finally, the contribution of taxonomic characteristics across studies indicates that they may also support early word learning. However, the fact that taxonomic effects were not significant in study 1 models that included word frequency in child directed speech, makes it unclear whether these findings are driven by learners' abilities or environmental input.

We next consider some limitations of our findings. First, the current project models normative lexical growth patterns for English nouns, which can capture broad, general patterns of lexical development. However, children show tremendous individual variability in their early word learning patterns, which may obscure individual differences. For example, prior findings suggest that there can be significant differences in the structural characteristics of networks of children at each end of the vocabulary size spectrum (e.g., Beckage et al., 2011). The vocabularies of children learning different first languages also have notable differences in structure, for example in regard to the proportion of early productive vocabulary made up of concrete nouns (Smith, Colunga, & Yoshida, 2003). Furthermore, although the MBCDI:WS covers a substantive proportion of children's early productive vocabularies, our use of the measure in combination with feature norms limited to concrete nouns ensured that all of our analyses covered only a subset of possible word knowledge. Future work is necessary to explore whether and how such individual differences, cross-linguistic differences, and more fine-grained measures of word knowledge that include other word classes might impact the normative models developed here. Second, the current findings are ultimately correlational in nature, necessitating further empirical work to test to our modeled results. For example, semantic priming studies could be used to differentiate between the "early taxonomic encoding," perceptual feature based taxonomic relational structure, and environmental explanations of the relation between AoA and taxonomic featural characteristics. Likewise, word learning studies could help determine whether higher-level structural characteristics play a direct role in learning or instead are a by-product of local, word and/or link level characteristics driving learning. We plan to pursue these questions in future work. Third, while our exploratory analyses in study 1 investigated possible influences of word frequency on our results, there are many other potential influences that we did not explore, including other ways of characterizing linguistic environments, such as via metrics of contextual diversity (e.g., Johns, Dye, & Jones, 2016), or other aspects of words that matter for learning that are not necessarily related to meaning, such as phonological characteristics (e.g., Stella et al., 2017; Swingley & Aslin, 2007). A

final limitation of the current study is that our feature norms are generated by adults, and therefore do not explicitly model developmental change in conceptual representations. Adults and toddlers are likely to encode/activate different aspects of familiar concepts, but our models do not explicitly account for this potential developmental variation. Future work is needed to develop detailed age-based estimations of semantic feature knowledge. In other work, we have attempted to gain insight into this process via another track – by asking adults to estimate whether toddlers are likely to understand specific semantic features (Peters et al., in prep). These ratings reveal that adults judge perceptual features to be more likely to be understood by toddlers. When we include the toddler-accessibility norms into our models we find that, while toddler accessibility accounts for a significant portion of the variance in AoA, perceptual feature characteristics remain significantly related to AoA for all levels of analysis (see supplementary analyses S-3 for details). These converging findings increase our confidence in our conclusion that perceptual features may be particularly important in developing early lexical representations.

In conclusion, the current project provides converging evidence that perceptual and, to a lesser extent, taxonomic properties of word meanings, associate with early lexical development, and may support early word learning and the development of early lexico-semantic networks from 16–30 months. This pattern lends support to accounts that posit that perceptual features of objects support early word learning, and only after the lexico-semantic system develops sufficiently does the apprehension of more complex/abstract features and relations become possible.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

5. References

- Arias-Trejo N, & Plunkett K (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3633–3647. [PubMed: 19933137]
- Arias-Trejo N, & Plunkett K (2013). What’s in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, 128(2), 214–227. [PubMed: 23688648]
- Barr DJ, Jackson L, & Phillips I (2014). Using a voice to put a name to a face: the psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General*, 143(1), 404. [PubMed: 23398179]
- Barsalou LW (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1435), 1177–1187. [PubMed: 12903648]
- Beckage N, Smith L, & Hills T (2011). Small worlds and semantic network growth in typical and late talkers. *PloS one*, 6(5), e19348. [PubMed: 21589924]
- Borovsky A, Ellis EM, Evans JL, & Elman JL (2016). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental science*, 19(6), 918–932. [PubMed: 26452444]
- Borovsky A, & Peters RE (under review). Vocabulary size and structure affects real-time lexical recognition in 18-month-olds.
- Braginsky M, Yurovsky D, Marchman VA, & Frank MC (2016). From uh-oh to tomorrow Predicting age of acquisition for early words across languages. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.

- Brysbaert M, & New B (2009). Moving beyond Kuera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977–990. [PubMed: 19897807]
- Brysbaert M, New B, & Keuleers E (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior research methods*, 44(4), 991–997. [PubMed: 22396136]
- Carey S (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Caselli NK, & Pyers JE (2017). The Road to Language Learning Is Not Entirely Iconic: Iconicity, Neighborhood Density, and Frequency Facilitate Acquisition of Sign Language. *Psychological Science*, 0956797617700498.
- Collins AM, & Quillian MR (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, 8(2), 240–247.
- Costa LDF, Rodrigues FA, Travieso G, & Villas Boas PR (2007). Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1), 167–242.
- Cree GS, McNorgan C, & McRae K (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 643.
- Cree GS, & 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(2), 163. [PubMed: 12825636]
- Dale PS, & Fenson L (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28(1), 125–127.
- Devereux BJ, Tyler LK, Geertzen J, & Randall B (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior research methods*, 46(4), 1119–1127. [PubMed: 24356992]
- Engelthaler T, & Hills TT (2017). Feature biases in early word learning: network distinctiveness predicts age of acquisition. *Cognitive science*, 41, 120–140. [PubMed: 26923664]
- Erdős P, & Rényi A (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6, 290–297.
- Fenson L, Bates E, Dale PS, Hartung JP, Pethick SJ, Reilly JS, ... & Thal DJ. (1996). *The MacArthur communicative development inventories*. Singular publishing Group, Incorporated.
- Frank MC, Braginsky M, Yurovsky D, & Marchman VA (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of child language*, 1–18.
- Freeman LC (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215–239.
- Friedman J, Hastie T, & Tibshirani R (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1. [PubMed: 20808728]
- Friedrich M, & Friederici AD (2006). Early N400 development and later language acquisition. *Psychophysiology*, 43(1), 1–12. [PubMed: 16629680]
- Gelman R (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive science*, 14(1), 79–106.
- Gillhooly KJ, & Logie RH (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, 12(4), 395–427.
- Gopnik A, & Meltzoff A (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child development*, 1523–1531.
- Griffiths TL, Steyvers M, & Firl A (2007). Google and the mind: Predicting fluency with PageRank. *Psychological Science*, 18(12), 1069–1076. [PubMed: 18031414]
- Grondin R, Lupker SJ, & McRae K (2009). Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory and Language*, 60(1), 1–19. [PubMed: 20046224]
- Groppe DM, Urbach TP, & Kutas M (2011a). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. [PubMed: 21895683]

- Groppe DM, Urbach TP, & Kutas M (2011b). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*, 48(12), 1726–1737. [PubMed: 21895684]
- Hollich G, Hirsh-Pasek K, & Golinkoff RM With Hennon E, Chung HL, Rocroi C, Brand RJ, & Brown E. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65, (3, Serial No. 262).
- Hills T (2013). The company that words keep: comparing the statistical structure of child-versus adult-directed language. *Journal of child language*, 40(3), 586–604. [PubMed: 22584041]
- Hills TT, Maouene M, Maouene J, Sheya A, & Smith L (2009a). Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition?. *Psychological Science*, 20(6), 729–739. [PubMed: 19470123]
- Hills TT, Maouene M, Maouene J, Sheya A, & Smith L (2009b). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, 112(3), 381–396. [PubMed: 19576579]
- Hills TT, Maouene J, Riordan B, & Smith LB (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of memory and language*, 63(3), 259–273. [PubMed: 20835374]
- Humphries MD, & Gurney K (2008). Network ‘small-world-ness’: a quantitative method for determining canonical network equivalence. *PloS one*, 3(4), e0002051. [PubMed: 18446219]
- Inhelder B, & Piaget J (1964). *The growth of logic in the child*. Routledge & Paul.
- Johns BT, Dye M, & Jones MN (2016). The influence of contextual diversity on word learning. *Psychonomic bulletin & review*, 23(4), 1214–1220. [PubMed: 26597891]
- Keil FC (1989). *Concepts, kinds and development*. Cambridge, MA: MIT Press.
- Landau B, Smith LB, & Jones SS (1988). The importance of shape in early lexical learning. *Cognitive development*, 3(3), 299–321.
- Landauer TK, & Dumais ST (1997). A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, 104(2), 211.
- Madole KL, Oakes LM, & Cohen LB (1993). Developmental changes in infants’ attention to function and form-function correlations. *Cognitive Development*, 8(2), 189–209.
- Mandler JM, & McDonough L (1996). Drinking and driving don’t mix: Inductive generalization in infancy. *Cognition*, 59(3), 307–335. [PubMed: 8706380]
- Maris E, & Oostenveld R (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. [PubMed: 17517438]
- McMurray B, Horst JS, & Samuelson LK (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological review*, 119(4), 831. [PubMed: 23088341]
- McRae K, Cree GS, Seidenberg MS, & McNorgan C (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior research methods*, 37(4), 547–559. [PubMed: 16629288]
- McRae K, Cree GS, Westmacott R, & De Sa VR (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 53(4), 360. [PubMed: 10646207]
- McRae K, De Sa VR, & Seidenberg MS (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126(2), 99. [PubMed: 9163932]
- Murphy GL, & Medin DL (1985). The role of theories in conceptual coherence. *Psychological Review*, 92(3), 289–316. [PubMed: 4023146]
- Newman ME, Moore C, & Watts DJ (2000). Mean-field solution of the small-world network model. *Physical Review Letters*, 84(14), 3201. [PubMed: 11019047]
- Oakes LM, & Rakison DH (2003). Issues in the early development of concepts and categories In Rakison & Oakes (Eds.) *Early category and concept development*, 3–23.
- Perry LK, Perlman M, & Lupyan G (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PloS one*, 10(9), e0137147. [PubMed: 26340349]
- Perry LK, Perlman M, Winter B, Massaro DW, & Lupyan G (2017). Iconicity in the speech of children and adults. *Developmental Science*.

- Peters R, McRae K, & Borovsky A (in prep). Semantic Feature Production Norms and Toddler-Accessibility Ratings for Early-Learned Noun Concepts: Extending McRae et al.'s (2005) Semantic Feature Norms.
- Pexman PM, Lupker SJ, & Hino Y (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review*, 9(3), 542–549. [PubMed: 12412895]
- Pexman PM, Hargreaves IS, Siakaluk PD, Bodner GE, & Pope J (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, 15(1), 161–167. [PubMed: 18605497]
- Plaut DC, & Shallice T (1993). Perseverative and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of cognitive neuroscience*, 5(1), 89–117. [PubMed: 23972122]
- Pruden SM, Hirsh- Pasek K, Golinkoff RM, & Hennon EA (2006). The birth of words: Ten- month-olds learn words through perceptual salience. *Child development*, 77(2), 266–280. [PubMed: 16611171]
- Quinn PC, & Eimas PD (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of experimental child psychology*, 63(1), 189–211. [PubMed: 8812045]
- Quinn PC, & Eimas PD (1997). A reexamination of the perceptual-to-conceptual shift in mental representations. *Review of General Psychology*, 1(3), 271.
- Quinn PC, & Eimas PD (2000). The emergence of category representations during infancy: Are separate perceptual and conceptual processes required?. *Journal of Cognition and development*, 1(1), 55–61.
- Rabovsky M, & McRae K (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, 132(1), 68–89. [PubMed: 24762924]
- Rabovsky M, Sommer W, & Abdel Rahman R (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, 6, 11. [PubMed: 22347855]
- Rakison DH (2003). Parts, motion, and the development of the animate–inanimate distinction in infancy. *Early category and concept development*, 159–192.
- Rakison DH (2005). Developing knowledge of objects' motion properties in infancy. *Cognition*, 96(3), 183–214. [PubMed: 15996558]
- Rakison DH, & Cohen LB (1999). Infants' Use of Functional Parts in Basic-like Categorization. *Developmental Science*, 2(4), 423–431.
- Rakison DH, Lupyan G, Oakes LM, & Walker-Andrews AS (2008). Developing object concepts in infancy: An associative learning perspective. *Monographs of the Society for Research in Child Development*, i–127.
- Rochat Y (2009). Closeness centrality extended to unconnected graphs: The harmonic centrality index. In ASNA (No. EPFL-CONF-200525).
- Rogers TT, & McClelland JL (2004). *Semantic cognition: A parallel distributed processing approach*. MIT press.
- Rosch E, & Mervis CB (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*, 7(4), 573–605.
- Rumelhart DE (1990, 1). Brain style computation: Learning and generalization In *An introduction to neural and electronic networks* (pp. 405–420). Academic Press Professional, Inc..
- Samuelson LK, & Smith LB (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child development*, 69(1), 94–104. [PubMed: 9499560]
- Samuelson LK, & Smith LB (1999). Early noun vocabularies: do ontology, category structure and syntax correspond?. *Cognition*, 73(1), 1–33. [PubMed: 10536222]
- Schneider R, Yurovsky D, & Frank M (2015). Large-scale investigations of variability in children's first words. In *CogSci* (pp. 2110–2115).
- Sheya A, & Smith LB (2006). Perceptual features and the development of conceptual knowledge. *Journal of Cognition and Development*, 7(4), 455–476.

- Smiley SS, & Brown AL (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology*, 28(2), 249–257.
- Smith EE, & Medin DL (1981). *Categories and concepts* (p. 89). Cambridge, MA: Harvard University Press.
- Smith EE, Shoben EJ, & Rips LJ (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological review*, 81(3), 214.
- Smith LB (2000). Learning how to learn words: An associative crane. *Becoming a word learner: A debate on lexical acquisition*, 51–80.
- Smith LB, Colunga E, & Yoshida H (2003). Making an ontology: Cross-linguistic evidence. In Rakison D & Oakes L (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 275–302). London: Oxford University Press.
- Smith LB, & Heise D (1992). Perceptual similarity and conceptual structure. *Advances in psychology*, 93, 233–272.
- Smith LB, Jones SS, & Landau B (1996). Naming in young children: A dumb attentional mechanism?. *Cognition*, 60(2), 143–171. [PubMed: 8811743]
- Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, & Samuelson L (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, 13(1), 13–19. [PubMed: 11892773]
- Spelke E (1994). Initial knowledge: Six suggestions. *Cognition*, 50(1), 431–445. [PubMed: 8039373]
- Stella M, Beckage NM, & Brede M (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports*, 7, 46730. [PubMed: 28436476]
- Steyvers M, & Tenenbaum JB (2005). The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive science*, 29(1), 41–78. [PubMed: 21702767]
- Storkel HL (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of child language*, 36(2), 291–321. [PubMed: 18761757]
- Storkel HL, & Adlof SM (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language, and Hearing Research*, 52(2), 306–320.
- Swingle D, & Aslin RN (2007). Lexical competition in young children's word learning. *Cognitive psychology*, 54(2), 99–132. [PubMed: 17054932]
- Unger L, Fisher AV, Nugent R, Ventura SL, & MacLellan CJ (2016). Developmental changes in semantic knowledge organization. *Journal of experimental child psychology*, 146, 202–222. [PubMed: 26974015]
- Vigliocco G, Vinson DP, Lewis W, & Garrett MF (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive psychology*, 48(4), 422–488. [PubMed: 15099798]
- Vitevitch MS (2008). What can graph theory tell us about word learning and lexical retrieval?. *Journal of Speech, Language, and Hearing Research*, 51(2), 408–422.
- Vitevitch MS, Ercal G, & Adagarla B (2011). Simulating retrieval from a highly clustered network: Implications for spoken word recognition. *Frontiers in psychology*, 2, 369. [PubMed: 22174705]
- Von Holzen K, & Mani N (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, 113(4), 569–586. [PubMed: 22980955]
- Watts DJ, & Strogatz SH (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440–442. [PubMed: 9623998]
- Willits JA, Wojcik EH, Seidenberg MS, & Saffran JR (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: evidence from auditory priming. *Infancy*, 18(6), 1053–1075.
- Younger B (1990). Infants' detection of correlations among feature categories. *Child Development*, 61(3), 614–620. [PubMed: 2364738]
- Yurovsky D, Bion R, Smith L, & Fernald A (2012). Mutual Exclusivity and Vocabulary Development. In *Proceedings of the Cognitive Science Society* (Vol. 34, No. 34).

Figure 1.
A noun-feature network containing five neighbors of SPOON (Left) and a subset of the features of SPOON.

Figure 2.

First-15-words perceptual network: A lexico-semantic network of the first 15 words learned by an average child. Links are defined using only perceptual features. Node size and color are weighted by the degree (Left), clustering coefficient (Center), and harmonic centrality (Right) of the word. CC = Clustering coefficient. HC = Harmonic centrality.

Figure 3.

(A) Normative link count trajectories and (B) mean normalized link weight trajectories, for each of the four feature types, all as a function of Wordbank-AoA rank, in bins of 10 words.

Figure 4.

Normative trajectories of link count by rank of Wordbank-AoA for each of the feature types, along with their respective random-AoA trajectories, in bins of 10 words. *Note:* The charts are on *different* scales to facilitate comparison between random and normative networks.

Figure 5.

Normative trajectories of density by rank of Wordbank-AoA for each of the feature types, along with their respective Random-AoA trajectories, in bins of 10 words. *Note.* The charts are on the *same* scale.

Figure 6.

Normative trajectories of global clustering coefficient by rank of Wordbank-AoA for each of the feature types, along with their respective Random-AoA trajectories, in bins of 10 words.

Note. The charts are on the *same* scale.

Figure 7.

Normative trajectories of mean distance by rank of Wordbank-AoA for each of the feature types, along with their respective Random-AoA trajectories, in bins of 10 words. *Note.* The charts are on the *same* scale.

Figure 8.

Normative trajectories of the metric of small-world-ness by rank of Wordbank-AoA for each of the feature types, along with their respective Random-AoA trajectories, in bins of 10 words. *Note.* The charts are on *different* scales to facilitate comparison between random and normative networks.

Table 1.

Results for base local (word)-level multivariate model, and models including Child Directed Speech word frequency (+CDS Model), Adult Directed Speech word frequency (+ADS model), and both Child and Adult Directed Speech word frequencies (Full Model).

Variable	Base Model		+CDS	+ADS	Full Model	
	Coef.	95% CI	Coef.	Coef.	Coef.	95% CI
Constant	0.00	[-0.10, 0.10]	0.00	-0.01	0.04	[-0.03, 0.12]
NoF:						
Enc.	-0.04	[-0.16, 0.07]	-0.03	-0.07	0.01	[-0.08, 0.09]
Func.	-0.07	[-0.19, 0.04]	-0.07	-0.13 *	-0.04	[-0.13, 0.04]
Percep.	-0.26 ***	[-0.36, -0.15]	-0.14 ***	-0.29 ***	-0.10 *	[-0.18, -0.02]
Tax.	-0.16 **	[-0.26, -0.05]	-0.06	-0.11 *	-0.05	[-0.13, 0.03]
Freq.:						
CDS	—	—	-0.62 ***	—	-0.85 ***	[-0.96, -0.74]
ADS	—	—	—	-0.35 ***	0.16 **	[0.06, 0.26]
<i>F</i>	7.60		57.45	16.12	59	
<i>R</i> ² _{adj}	.07		.44	.18	.50	

Note. N=359. CI = confidence interval. Coefficients are β_{std} .

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 2.

Results for base neighborhood-level multivariate model, and models including Child Directed Speech word frequency (+CDS Model), Adult Directed Speech word frequency (+ADS model), and both Child and Adult Directed Speech word frequencies (Full Model).

Variable	Base Model		+CDS	+ADS	Full Model	
	Coef.	95% CI	Coef.	Coef.	Coef.	95% CI
Constant	0.00	[-0.10, 0.10]	0.00	0.00	0.00	[-0.08, 0.08]
Degree:						
Enc.	-0.06	[-0.05, 0.16]	0.06	0.03	0.08 *	[0.00, 0.16]
Func.	0.02	[-0.09, 0.12]	-0.09 *	-0.05	-0.06	[-0.14, 0.02]
Percep.	-0.15 *	[-0.27, -0.03]	-0.11 *	-0.22 ***	-0.07	[-0.15, 0.02]
Tax.	-0.10	[-0.21, 0.01]	-.02	-0.09	-0.00	[-0.08, 0.08]
CC:						
Enc.	-0.01	[-0.10, 0.11]	-0.00	-0.02	-0.01	[-0.08, 0.07]
Func.	-0.06	[-0.17, 0.04]	-0.03	-0.05	-0.05	[-0.12, 0.03]
Percep.	-0.04	[-0.15, 0.08]	-0.08	-0.08	-0.07	[-0.15, 0.01]
Tax.	0.09	[-0.02, 0.19]	.07	0.04	0.01	[-0.08, 0.09]
Freq.:						
CDS	—	—	-0.66 ***	—	-0.88 ***	[-0.99, -0.77]
ADS	—	—	—	-0.39 ***	0.17 **	[0.06, 0.27]
<i>F</i>	2.221		32.75	7.94	35.96	
<i>R</i> ² _{adj}	.026		.44	.15	.50	

Note. N=359. CC = Clustering Coefficient. CI = Confidence Interval. Coefficients are β_{std} .

*
 $p < .05$.

**
 $p < .01$.

 $p < .001$.

Table 3.

Results for base lexicon-level multivariate model, and models including Child Directed Speech word frequency (+CDS Model), Adult Directed Speech word frequency (+ADS model), and both Child and Adult Directed Speech word frequencies (Full Model).

Variable	Base Model		+CDS	+ADS	Full Model	
	Coef.	95% CI	Coef.	Coef.	Coef.	95% CI
Constant	0.00	[-0.10, 0.10]	0.00	-0.02	0.04	[-0.03, 0.12]
HC:						
Enc.	-0.02	[-0.12, 0.09]	0.02	-0.05	0.06	[-0.02, 0.13]
Func.	0.02	[-0.09, 0.13]	-0.08 *	-0.07	-0.06	[-0.53, 0.02]
Percep.	-0.15 *	[-0.26, -0.03]	-0.09 *	-0.22 ***	-0.04	[-0.13, 0.04]
Tax.	-0.11 *	[-0.22, -0.01]	-0.05	-0.10	-0.02	[-0.10, 0.06]
Freq.:						
CDS	—	—	-0.66 ***	—	-0.88 ***	[-0.99, -0.76]
ADS	—	—	—	-0.40 ***	0.16 **	[0.05, 0.27]
<i>F</i>	3.42		57.68	14.98	58.72	
<i>R</i> ² adj	.026		.44	.17	.50	

Note. N=359. HC = Harmonic Centrality. CI = confidence interval. Coefficients are β_{std} .

* $p < .05$.

** $p < .01$.

*** $p < .001$.