Title
Contextual Diversity and the Associative Structure of Adult Language in Early Word Learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Hills, Thomas
Maouene, Josita
Riordan, Brian
et al.

Publication Date
2009

Peer reviewed

Contextual Diversity and the Associative Structure of Adult Language in Early Word Learning

Thomas T. Hills (**thomas.hills@unibas.ch**)
University of Basel, Cognitive and Decision Sciences, Basel, Switzerland

Josita Maouene (jcmaoune@indiana.edu), **Brian Riordan** (**briordan@indiana.edu**),
Linda B. Smith (smith4@indiana.edu)
Department of Psychological and Brain Sciences, 1101 E. Tenth Street
Bloomington, IN, 47405 USA

Abstract

Previous studies have demonstrated that the statistical properties of adult generated free associates can predict the order of early word learning in children. In this paper, we investigate the cause of this phenomenon. We propose that early word learning may be driven by the contextual diversity of words in child directed speech, which is in turn correlated with an underlying associative structure in adult language. We present evidence for this hypothesis by analysing the co-occurrence of words in the CHILDES corpus of child directed speech. We find that a word's contextual diversity—its number of unique neighbors—predicts the order of early word learning and is highly correlated with a word's associative diversity. Using longitudinal network analysis on developing early semantic networks from 15 to 30 months, we also find evidence for a specific growth process called preferential acquisition, in which words with more diversity in the learning environment are learned earlier than less diverse words. Only adjectives support preferential attachment—a process based on the structure of known words—and this is consistent with the evidence that adjective learning is a strongly facilitatory process, in which the learning of an adjective is enhanced by comparison with other similar adjectives.

Keywords: network analysis, age of acquisition, child directed speech, free association, contextual diversity, corpus analysis, nouns, adjectives

Introduction

Adult free-associations predict the order of early noun acquisition (Hills et al., in press; Hills et al., 2008; Steyvers & Tenenbaum, 2005). In a standard free association task, researchers provide cues and subjects provide the first word that comes to mind (the target). Words that are recalled as targets for a larger set of cues are acquired earlier in development than words that are the recalled target for a smaller set of cues. A clear question is why do adult associations predict age of acquisition (AoA)?

Children do not have direct access to information about adult free associates, therefore the predictive power of associates must be due to their correlation with some other property of the learning environment. One explanation for the success of associates is that they represent a proxy for structural information contained in adult language. The words children acquire earlier are the ones elicited by many different contexts in adult free association tasks; perhaps then, the key factor relevant to acquisition is contextual diversity. The words adults produce in many different contexts in the free association task may be the ones that they also produce in many different contexts in their language to children. In part, early-learned words may have more associative relations because they occur in more contextually diverse contexts in adult language. If this idea is correct, then we should see a similar associative structure in child directed speech to that observed in standard free-association tasks and, moreover, we should also find contextual diversity—measured by number of unique contexts—is a predictor of age of acquisition.

Contextual diversity has not been directly studied in early vocabulary development, but a number of studies suggest its importance. Work by Hayes and Clark (1970) found that adults, when listening to an artificial speech stream without clear word boundaries, detected word boundaries in proportion to the transitional probabilities between phonemes—with words in more diverse phonemic backgrounds being more easily differentiated. Saffran et al. (1996a, 1996b) also demonstrated this result for adults, and went on to show that eight-month-olds could make similar word non-word distinctions based on transition probabilities when presented with unparsed speech stream, e.g., identifying "*bidaku*" after listening to speech of the form "*bidakupadotigolabubidaku*."

This suggests that early lexical learning may be particularly sensitive to a lexical item's contextual diversity, because diversity provides critical information for differentiating word boundaries and potential word-object mappings. That is, items that appear in many varied contexts are more easily disambiguated, both with regards to the primary speech data and also with regards to separating object from background. If contextual diversity actually drives language acquisition, then this leads to the prediction that the age at which a word is acquired should be a consequence of its contextual diversity in child directed speech.

In this paper, we use the CHILDES corpus of child directed speech (MacWhinney, 2000) to investigate the contextual diversity in the structure of adult language. In particular, we ask, how does the contextual diversity in child directed speech correlate with the statistical properites of free associates. Secondly, we ask if contextual diversty might be a better predictor of age of acquisition than associates—with the implication that the predictive power of associates is explained by their relationship to contextual diversity in child directed speech.

A second question we address here is whether the predictive power of associates and contextual diversity also apply to non-noun word classes. There is reason to believe that they might not. Different word classes are learned at different times (e.g., Dale & Fenson, 1996), and possibly by different mechanisms. For example, Waxman & Klibanoff (2000) demonstrated that adjective learning was strongly facilitated by comparison either across adjectives or across basic level nouns—whereas children tend to default to basic level noun mapping when hearing novel words with novel objects (e.g., Waxman & Kosowski, 1990). Also, Tomasello's (2000) Verb Island hypothesis might be taken to suggest that verbs are learned differently from other words, taking on more independent representations in their earliest learning. We examine the predictive power of associations and contextual diversity for each word class.

Our main focus is therefore to examine the relationship between contextual diversity and the associative structure of child directed speech as a possible source of information relevant to learning those words, and secondly to examine how these structural cues may differ for different word classes.

Methods

The words. The words were taken from the MacArthur-BatesCommunicative Developmental Inventory or MCDI (Dale & Fenson, 1996), Toddler version. This inventory is a checklist of 574 words typically acquired by children learning English and normatively included in the productive vocabularies of 50% of children at 30 months of age. For our analyses, we excluded 42 words that were never recalled as target words in the free association norms, or were words about time. This left 532 words, of which 330 were nouns, 96 were action words, 58 were descriptive words, 21 were pronouns, and 48 were function words consisting of quantifiers, articles, helping verbs, and connecting words. We let the age of acquisition for a word equal the first month at which the word was produced by more than 50% of the children in the normative tables of the MCDI.

Associates. We used the adult-generated University of Southern Florida Free Association Norms (Nelson, et al., 1999). These were collected by providing subjects with a word (the cue) and asking them to provide the first word that came to mind (the target). This establishes a cue-target relationship, e.g., when provided with the cue word CAT many subjects provide the target word DOG. The norms consist of approximately 5000 cue words and their related targets.

CHILDES. In order to train our model of contextual diversity (see below), we used a 2 million word corpus of caregiver speech derived from the CHILDES database (Riordan & Jones, 2007). In this corpus, a large number of word forms were standardized to reduce orthographic variation introduced by varying transcription conventions. Words were also converted to stemmed forms (e.g. cats → cat) under the assumption that word meaning in child-directed speech varies little across inflectional variants. Words were aligned with their appropriate matching word in the MCDI data.

Co-occurrence. To generate a lexical semantic representation based on the co-occurence of words in CHILDES, we built matrix representations of word co-occurrences using a process similar to the Hyperspace Analogue to Language (HAL) (Lund & Burgess, 1996) and the word co-occurrence detector (Li et al, 2004). For a corpus of N unique words, we formed an NxN network, where each cell, ij, is filled according to the following rule: a moving window of size k moves word-wise through the corpus, with each word, i, adding to cell ij, a value equivalent to the window size minus the number of intervening words to the subsequent word j, for all words in the window. We also used three different window sizes of 5, 10, and 15. The results were not qualitatively different across window sizes, thus we only report results for a window size of 15. After converting all cell entries greater than one to one, the sum of columns and rows provides a straightforward indication of a word's contextual diversity. Frequency counts were taken as the number of occurrences of a given word. Our analysis only uses the 5000 most frequent words in the corpus.

The networks. To construct the developing networks used for the following analyses, we let words represent nodes. Links between words were formed depending on whether or not two words contained an associative or co-occurrence relation. For the association network, each word pair was connected by a directed link from the cue word to the target word if that cue-target relationship was reported in the association norms. For the co-occurrence networks, each word pair was connected if the two words co-occurred in child directed speech. We then created 15 networks, for each month between 15 and 30 months, by including only words acquired by that month. This generated a developmentally ordered set of 15 association and 15 co-occurrence networks. For the 30 month networks used for preferential acquisition, we only use the words in the given word class. For the adult network, we use the combined 532 words from all word classes.

Results

As reported in the introduction, previous research established the relationship between age of acquisition for nouns and its associative relationships in the adult free association norms. The number of associative relationships was taken as the count of the number of distinct cue words for which the target word was recalled. For clarity, we will call this value the *associative indegree*. To investigate the relationship between associative indegree and age of acquisition for different word classes, we investigate the predictive power of associative indegree (in the 30 month network) on age of acquisition in Table 1. These results also present the R^2 contribution of associative indegree and word frequency, each after controlling for the other.

Table 1: Effects of log-transformed associative indegree (Log-AI) and log-transformed word frequency (Log-WF) on age of acquisition.

Word class	Effect on AoA in (ΔR2 in %) log-AI	log-WF	Log-AI after Log-WF	Log-WF after Log-AI	Log-WF + Log-AI
ALL	10.1***	3.1***	8.0***	1.0*	11.1***
Nouns	7.9***	39.6***	1.0*	32.8***	40.6***
Verbs	14.7***	12.2***	5.5*	3.1ns	17.7***
Adjectives	13.6**	13.8**	3.9ns	4.1ns	17.7**
Function	20.3***	20.9***	17.8***	18.2***	38.7***

The results from Table 1 indicate that associative indegree does make a significant contribution to age of acquisition, and that for nouns, verbs, and function words, this contribution is above and beyond that contributed by word frequency alone. This confirms and extends our previous findings (Hills et al., in press), demonstrating that associative indegree is effective for a broader class of nouns, as well as other word classes. In all cases, the sign of the coefficient for log-AI was negative, indicating that words that are the target for a larger set of cues (have more associates) are more likely to be learned earlier than words associated with a smaller set of cues.

These results are potentially a consequence of a statistical learning strategy based on contextual diversity. One way to approach this hypothesis is to ask what aspects of the structure of child directed language are similar to adult generated free associates, and are these aspects also predictive of age of acquisition. If free associates are simply dependent variables that are themselves the products of cognitive semantic knowledge representations, then the structure of child-directed speech should actually be more informative with respect to age of acquisition than associates. To investigate this possibility, we took the CHILDES degree (CHd)—using the sum of both column and row for a given word—as an independent variable, and controlling for frequency of the word in the CHILDES corpus, we regressed these on the AoA from the MCDI for each word in a given word class (see Table 2). Thus, in part, Table 2 parallels Table 1, with CHd replacing AI.

First, note that CHd is a predictor of AoA, and that this is true even after controlling for word frequency. Second, the overall fit of a word's diversity in CHILDES and its frequency is as good or better than associative indegree as a predictor of age of acquistion. This is true of all word classes except function words. Log-transformed nouns are possibly another case where this may not be true, but nouns are also the least frequent (and the least diverse) word class, and may therefore lose information in the log transformation. The sign of the coefficients are always negative, with more contextually diverse nouns being learned at younger ages. Finally, note that the correlation between a word's diversity in CHILDES is also highly correlated with its associative indegree. These results strongly suggest that the performance of adult associates is due to their correlation with the structure of child directed speech—in particular, their correlation with a word's contextual diversity.

Table 2: Results of log-transformed CHILDES degree (Log-CHd) on age of acquisition. Items in parenthesis are not log-transformed for CHd.

Word class	Effect on AoA in (ΔR2 in %) Log-CHd	Log-CHd after Log-WF	Log-CHd + Log-WF	Correlation Log-CHd & Log-AI
ALL	2.1*** (0.0ns)	4.2***(13.6****)	7.3*** (16.7***)	27.3*** (4.2ns)
Nouns	36.8*** (25.7***)	0.1 (2.2***)	40.4*** (41.9***)	61.8*** (55.1***)
Verbs	9.5** (5.1*)	7.2** (5.7*)	19.5*** (18.0***)	55.8***(36.2***)
Adjectives	10.3** (8.8*)	13.0** (1.4ns)	26.8** (15.2*)	52.9***(39.1**)
Function	10.8*** (19.1***)	11.2** (0.1ns)	32.1*** (21.1**)	-7.2ns(7.9ns)

We also examined how much of the variance is explained by CHd after controlling for associates. If our initial hypothesis is true, and contextual diversity in CHILDES is the force driving the age of acquisition effect found for associates, then most of the variance explained by associative indegree should disappear if we first take out the contribution made by CHILDES degree. Table 3 shows that, for any given word class, CHILDES degree explains significant variance after taking out that explained by associative indegree. Associative indegree is still predictive, but to a lesser extent, explaining less of the variance for both nouns and function words. Associative indegree does retain a high overall explanatory power when all words are combined.

Table 3: Results of log-transformed CHILDES degree (log-CHd) and Log-transformed associative indegree in explaining age of acquisition, after controlling for other factors. Items in parenthesis are not log-transformed for CHd.

Word class	Effect on AoA in (ΔR2 in %) Log-CHd after Log-WF and Log-AI	Log-AI after Log-WF and Log-CHd
ALL	6.9***(6.5***)	13.7***(3.0***)
Nouns	0.7 (2.3***)	0.0 (0.0)
Verbs	10.9***(3.2)	10.7***(4.5*)
Adjectives	11.3**(0.0)	5.0(5.4)
Function	5.4*(0.0)	4.7(10.4*)

In summary, the above analyses suggest that contextual diversity, as measured by the count of a words co-occurrence with unique words in child directed speech, is a significant predictor of age of acquisition for all word classes. The results also indicate a significant relationship with associative indegree, but do not completely remove the possibility of an independent effect of associates—one that is not related to contextual diversity.

In prior work, beyond establishing the efficacy of associates to predict AoA, we also hypothesized and found support for a specific growth process called *preferential acquisition.* This growth process is consistent with the contextual diversity hypothesis because it proposes that words are learned in relation to their contextual diversity in the learning environment. In prior work, we only examined diversity as measured by a nouns associative indegree. Here we extend this analysis to CHILDES indegree, examing the three growth hypotheses we investigated in Hills et al. (In press).

The three growth rules are as follows (Figure 1): preferential attachment – based on the connectivity of known words; preferential acquisition – based on the connectivity of new words to all words in the learning environment; and a third intermediate possibility, lure of associates, based on the connectivity of new words to known words. For preferential attachment, a word is more likely to be learned if it attaches to an existing already *known* word in the network that is itself well attached. In this way, richly connected words become more richly connected. In contrast, with preferential acquisition, a word is more likely to be learned if it is attached to many other words *in the learning environment*. The lure of associates lies between the above two; at the time-of-acquisition, the child learns next the word that attaches to the most already known words.

We asked which model for growth best fits the CHILDES degree using a maximum likelihood test. Our basic growth model determines how strongly the growing network weighs the value of new words at each month in the growth of the network, with value determined by the model. We do this using a parameter, β, which we fit to an exponential ratio of strengths rule. We calculate the probability that a word, w_i, is added to the network at a given month based on its value, d_i, using the following:

$$P(w_i) = \frac{e^{\beta d_i}}{\sum_j e^{\beta d_j}}$$

Here, β represents the sensitivity of the acquisition process to d_i. In particular, positive values of β mean that words with higher values of d_i are more likely to be acquired early, whereas negative values of β mean that words with low values of d_i are more likely to be acquired early. A β value of exactly zero would indicate that order of acquisition is not sensitive to d_i. We let d_i represent the degree values ("Value" in Figure 1) for each word calculated with respect to each model. For example, with the lure of associates model, d_i is equivalent to the in-degree of the word i at the time of acquisition. The denominator is calculated for all words that are not yet learned at the start of the month for which the word in the numerator is acquired. The log of the $P(w_i)$ values, for each acquired word, is then added to produce the log likelihood.

$$-\log(L(\beta)) = -\sum_i \log(P(w_i))$$

We then find the β^* that minimizes the above log likelihood function using a standard optimization procedure.

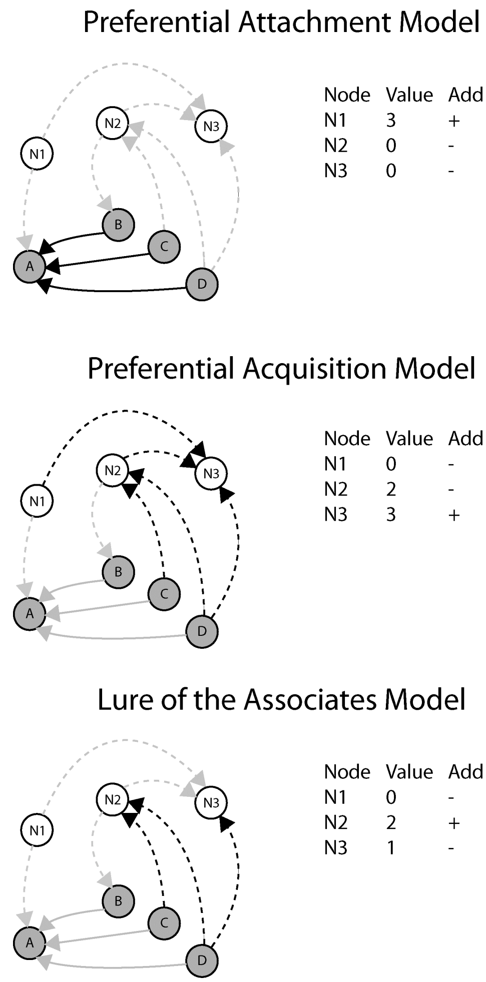

Figure 1.The three growth models for a simplified network. Each of the networks is the same, but the growth models differently value each unknown word (possible new words are shown in white, existing nodes are represented in grey; links relevant to the growth model are shown in black, unimportant links in grey). "Add" indicates which node is favored for learning by the growth model. With Preferential Attachment, the value of the new node is the average degree of the known nodes it would attach to. With Preferential Acquisition, the value of the new node is its degree in the full network. With the Lure of the Associates, the value of new node is its degree with respect to known words. This figure is taken from Hills et al. (In press).

Table 4 presents the results for the model comparison, using CHILDES degree. The results indicate that for most word classes, preferential acquisition or lure of the associates are the best fitting models. As shown in Figure 1, these models only differ by the inclusion of unknown words when moving from lure of the associates to preferential acquisition. That the best fit for function words is preferential acquisition based on the adult network is most likely due to the fact that function words gain most of their co-occurrence information from non-function words, which—for this word class—are only present in the adult network. Finally, adjectives are the only word class that supports preferential attachment.

Table 4: Results of the log likelihood fits for the three growth models. Model fits are ranked by their log likelihood value, with 1 representing the best fit. Models that were not different from random by AIC are represented with an 'r'. We take a conservative estimate of 6 G^2 units as our measure of significance, and do not rank models that are more distant than 6 units from the best model. P. Attach. = preferential attachment. P.Acq = Preferential acquisition using the designated network (30 months or Adult).

	All	Noun	Verb	Adj	Fun
P. Attach.	r	r	r	1	r
Lure of Assoc.	1	1	2	r	-
P. Acq-30mts	2	2	1	r	2
P. Acq-Adult	r	-	3	r	1

Discussion

This paper provides a first look at how the statistical structure of child directed speech may facilitate a specific pattern of early word learning. Our results provide a possible explanation for the success of associates in predicting the order of early word learning. We do this by demonstrating that the associative indegree of a word is strongly correlated with the contextual diversity of a word in child directed speech. We further show that both associative indegree and the number of co-occurring neighbors of a word in child directed speech are predictors of the order of acquisition for words in the first 30 months.

The role of contextual diversity in learning is gaining momentum and we feel this explanation for the success of associates in prior work is consistent with literature on early language learning. It is also consistent with a growing literature on contextual diversity in lexical decision times (Adelman et al., 2006; Steyvers & Malmberg, 2003; Hicks et al., 2005; Recchia et al., 2008). A high level explanation for the effect of contextual diversity is provided by the rational analysis of memory (Anderson & Milson, 1989; Anderson & Schooler, 1991), which is based on the principle of likely need. This principle suggests that words in more contexts are more likely to be needed in any new context, and thus should be learned earlier. In the introduction, we provide instead an argument based on mechanism—high diversity words are more easily disambiguated; they are more readily isolated from the background and mapped to their corresponding referents.

Our results show that the predictive power of contextual diversity is effective for all word classes, but differentially so. Nouns show the strongest predictive effect for contextual diversity in child directed speech, but much of this effect disappears after controlling for word frequency. The necessary precedence of frequency, however, is far from established. In Table 1, frequency across all words is much less predictive than associative indegree. Frequency and contextual diversity are related, but word repitition is clearly not sufficient to drive early word learning—otherwise, function words would be learned earliest.

The final result is that preferential acquisition, and its close neighbor, lure of the associates, are still supported as the most favorable growth models for development in early semantic networks. However, this effect does not hold for all word classes. For descriptive words (adjectives), preferential attachment is the dominant model, significantly outperforming all other models. This result is consistent with the known differences between noun and adjective learning (Gasser & Smith, 1998; Sandhofer & Smith, 2007). In adjective learning, the role of comparison appears to play a much stronger role than in learning with other word classes. For example, the process of helping children to extend adjectives appropriately is strongly facilitated by presenting a comparison object that differs along the adjectival dimension, or that differs in everything but the adjectival dimension (Waxman & Klibanoff, 2000). Further, learning new adjectives appears to be strongly predicted by how many other adjectives one knows for the class of properties (e.g., Sandhofer & Smith, 1999; Backscheider & Shatz, 1993) In this way, strong knowledge of one adjective facilitates the learning of other related adjectives, creating a preferential attachment growth process.

We note that the age of acquisition effect—in which lexical decision times are related to the age of a word's acquisition—might be interpreted as a situation where early word learning is driving adult retrieval times (Ellis and Morrison, 1998). In this paper we explore an alternative possibility, in which the age of acquisition effect is a consequence of the associative structure of language, which in turn drives age of acquisition via preferential acquisition. In the final analysis, the direction of causation may well go in both directions—there is evidence in support of both (e.g., Recchia et al., 2008; Stewart & Ellis, 2008). We propose that this evidence is best explained by seeing word learning as a self-reinforcing dynamical system, in which the earliest learned words become more easily retrieved during speech, and thus reinforce the learning of these words earliest in future generations through a process involving contextual diversity.

Acknowledgments

This work was supported by the National Institute of Health, T32 Grant # HD 07475.

References

Adelman, J.S., Brown, G.D.A., Queseda, J. F. (2006). Contextual Diversity, Not Word Frequency, Determines Word-Naming and Lexical Decision Times. *Psychological Science* 17(9), 814-823.

Anderson, J. R. & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96(4), 703-719.

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science* 2(6), 396-408.

Backscheider, A. G. & Shatz, M. (1993). Children's acquisition of the lexical domain of color. In K. Beals et al. (Eds), *What we think, what we mean, and how we say it*, Vol. 2 (pp. 11–21). Chicago: The Chicago Linguistic Society.

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers,* 28, 125-127.

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 24, 515-523.

Gasser, M., & Smith, L. B. (1998). Learning nouns and adjectives: A connectionist account. *Language and Cognitive Processes,* 13, 269-306.

Hayes, J.R., & Clark, H.H. (1970). Experiments in the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the Development of Language*. New York: Wiley.

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). An observation on the role of contextual variability in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1160-1164.

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

Hills, T., Maouene, M., Maouene, J., Sheya, A. & Smith, L. (In press). Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychological Science.*

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114(1), 1–37

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavioral Research Methods, Instruments, and Computers,* 28(2), 203-208.

Li, P., Farkas, I., MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks* 17, 1345-1362.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Recchia, G., Johns, B. T., & Jones, M. N. (2008). Context repetition benefits are dependent on context redundancy. *Proceedings of the 30th Annual Cognitive Science Society, Washington, DC, USA, July 23-26.*

Riordan, B., and Michael N. J.. (2007). Comparing semantic space models using child-directed speech. In D. S. MacNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Pgs. 599-604.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Sandhoffer, C., & Smith, L. B. (2007). Learning adjectives in the real world: how learning nouns impedes learning adjectives. *Language Learning and Development*, 3(3), 233-267.

Smith, L., & Yu, C. (2007). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106, 1558-1568.

Stewart, N., & Ellis, A. W. (2008). Order of acquisition in learning perceptual categories: a laboratory analogue of the age of acquisition effect? *Psychonomic Bulletin and Review,* 15, 70-74.

Steyvers, M., & Malmberg, K. (2003). The effect of normative context variability on recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29(5), 760-766.

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

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

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, 36, 571-581.

Waxman, S. R., & Kosowksi, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development*, 61, 1461-1473.