Sound versus meaning: What matters most in early word learning?

Sarah Devi Sahni and Timothy T. Rogers

Abstract—**Previous work suggests that phonological neighborhood density is a key factor in shaping early lexical acquisition. Such studies have, however, have not considered how semantic neighborhoods may influence word-learning. We studied how phonological and semantic densities affect both comprehension and production of nouns from the Macarthur-Bates Communicative Development Inventory (MCDI). New measures of semantic and phonological densities, along with child-directed word frequency counts were used to predict the percentage of children who know each word at different ages (8 - 30 months) as indicated in MCDI lexical norms. Production was predicted by frequency and phonological density at all time points, replicating previous research. Semantic density predicted production only at 30 months. Comprehension norms were predicted by frequency and semantic density, and never by phonological density. Two- and three-way interactions reveal that semantic density may moderate effects in production, while sound density may moderate effects in comprehension.**

Index Terms—**cognitive development, lexical acquisition, semantic neighborhoods, phonological neighborhoods**

I. Introduction

As is now widely known, children are sensitive to statistical regularities in speech, and likely capitalize on these regularities when learning their native language. A central goal of current research in language acquisition is to understand precisely which statistical regularities are important and how this information is extracted and used to support different aspects of language acquisition and processing. In the domain of word-learning, the majority of research has focused on the role of phonological structure with the aim of understanding how such structure might be used to support segmentation of the speech stream, the addition of new words to the "mental lexicon," and the ability to map from lexical representations to the referents they denote. While this research program has been exceedingly productive in many respects, it has also raised some puzzling questions about how phonological structure influences word acquisition.

A variety of recent studies suggest that infants struggle to learn words from relatively *dense* phonological neighborhoods. That is, when a new word is similar in sound to words they already know, young children may have difficulty learning to map it to a new referent, despite being able to perceive that the new word differs from the known words. For example, using a habituation task, Werker and colleagues [1] found that 14-month-old infants were able to associate two novel labels with paired novel objects, but only when the novel words were phonologically distinct (*lif* vs *neem*) and not when they were phonologically similar (*bih* vs *dih*)—even though infants were capable of discriminating the similar-sounding words in a task that did not require them to associate the words with novel objects. Along similar lines, Swingley and Aslin [2] attempted to teach 18-month-olds novel neighbors of words already in the lexicon (e.g. *tog*, neighbor to *dog*, and *gall* neighbor to *ball*). Although infants can reliably distinguish these mispronunciations from their source words [3, 4], the infants were unable to map them to novel objects—again suggesting that phonological similarity inhibits word-learning. Together, this research suggests that high phonological density should slow the acquisition of new words.

In contrast, phonological density of real words appears to <u>aid</u> word learning. Storkel [5] examined whether phonological neighborhood density (together with word frequency and word length) could predict the age of acquisition of early vocabulary items from the Macarthur-Bates Communicative Development Inventory (MCDI) lexical production norms [6]. Perhaps surprisingly, she found that words from phonologically dense neighborhoods were acquired earlier than words from sparse neighborhoods, even after accounting for effects of word frequency and length. This effect also interacted with word frequency and length, being more pronounced for lower-frequency and shorter words. Supporting these results, an artificial word-learning study indirectly suggests that density may facilitate word-learning [7]. Graf Estes found that 18-month-old infants were able to learn nonsense labels that contain common sound sequences but not those containing rare sequences. Though the phonological neighborhood density of these labels was not controlled, the labels containing common sound sequences had high phonotactic probability and thus were phonotactically more similar to other labels in the lexicon compared to those containing rare sound sequences—again suggesting that density (in this case phonotactic density) facilitates word learning.

What accounts for these seemingly contradictory conclusions? In this paper, we consider two possible explanations. First, these studies have focused exclusively on phonological similarity structure, but have not considered how such effects might interact with *semantic* similarity structure. When a child learns a real word, she learns, not just what it sounds like, but also what it means. Perhaps the ease with which new words are learned is jointly influenced by both semantic and phonological similarity structure. For instance, in Werker's study, perhaps toddlers were unable to associate *bih* and *dih* to the different novel objects not *solely* because the two words had similar sounds, but also because the two novel objects may not have been semantically distinct to the infants. While the objects were visibly different, both were novel to the infant, they moved in similar paths, and there was no context for either object (i.e., there was no information about how it interacts with other objects or how people may use it). Moreover, there is evidence to suggest that additional experience with objects may enable infants to learn the separate labels for the two objects [8].Thus to understand early word learning, it may be important to consider the joint contributions of phonological and semantic similarity structure, and how these factors interact. That is the first goal of this paper.

Second, the studies of Werker [1] and Swingley and Aslin [2] employed *receptive* tasks to assess word learning. That is, the 14-18-month-olds were not required to produce the words they were learning. Infants were familiarized to label-object pairs, and learning was assessed via looking time measures. In contrast, Storkel's corpus analysis investigated the influence of phonological structure on lexical *production* norms. Is it possible that phonological similarity structure has different effects on production and comprehension? Answering this question is the second goal of this paper.

To achieve these goals, we undertook an analysis of the MCDI lexical norm database modeled on Storkel's work, but with two important differences. First, in addition to phonological density and word frequency, we included a measure of the *semantic density* as a predictor in the analyses. Second, we ran separate regression analyses for both the comprehension (Infant Knows) and production (Infant Says / Toddler Says) components of the database. These analyses allowed us to consider the joint contribution of both phonological and semantic similarity structure to both comprehension and production norms.

II. Methods

The outcome measures used in these analyses were lexical norms, drawn from the Lex2005 database [6], indicating the proportion of children that can understand or produce a given word at different ages. Words were drawn from the MCDI Infant Knows (8 to16 months), Infant Says (8 to 16 months), and Toddler Says (16 to 30 months) checklists.

Each word was assigned a phonological description based on the coding-scheme described by Joanisse & Seidenberg [9]. In this scheme, each phoneme of a word is represented with a set of binary features (see Appendix A) corresponding to phonetic properties of the utterance such as place of articulation, voicing, and so on. These phoneme-feature-vectors are strung together to construct a single vector representing the entire word. Two of the original 18 features were removed from the phonological feature set due to complete redundancy with another feature (specifically, voiceless was completely redundant with voiced and obstruent was completely redundant with sonorant). Words were centered on the first vowel so that rhyming words have similar representations (e.g. *doll, ball,* and *crawl* will only differ in the first two phonemes). Up to two consonants preceded the first vowel and a maximum of 12 phonemes were allowed per word. Words that did not have two consonants preceding the first vowel were padded with the appropriate number of leading zeroes. Similarly, words shorter than 12 phonemes had empty final slots filled with zeroes.

Each word was also assigned a semantic description based on the semantic feature-coding scheme developed by Howell, Jankowicz, & Becker [10]. In this scheme, each word is evaluated on 97 perceptually-grounded features to which infants and children are likely to have direct perceptual access. These include features such as size, color, external parts, patterns of movement, and so on (see Appendix A). Continuous feature dimensions such as size (which varies continuously from small to large) are coded as real-valued numbers ranging between 0 and 1, whereas binary features (e.g. "has paws") were coded with 0 or 1 values.

The corpus used for the analyses was composed of a subset of nouns from the MCDI checklist. Words were only included if both a semantic and sound representation could be created. That is, words that were not included in Howell et al.'s original set or that violated the template for the phonemic coding (i.e. had more than two consonants preceding the first vowel or contained more than 12 phonemes) were excluded from the corpus. For the infant data sets, this totaled 207 words. For the toddler data set there were 328 words.

Neighborhood densities for all words were computed in the same way for both the phonological and semantic representation vectors. Specifically, we computed the Euclidean distance D between a given item *i* and all other items *j* in the corpus, and calculated a density function as follows:

$$\sum_j \frac{1}{D_{ij}^2}$$

Intuitively, this sum will be large for an item with many close neighbors (where the squared distance for each is small), and will be small for items with few close neighbors (where D^2 for each is large). Thus the larger the value, the more dense the item's neighborhood. The influence of any given neighbor on this sum drops exponentially with distance. The virtue of such a measure is that it permits all words in the lexicon to have some influence on an item's density, rather than considering, for instance, just those local neighbors that differ from the target word in a single phoneme.

In addition to sound and semantic density, the frequency of each word in speech to children is likely to influence acquisition. Word frequency counts were acquired from several age-appropriate corpora within the CHILDES database [11]. Frequency of these words in the input to children 16 months and younger was calculated for analyses predicting infant norms. Frequencies were obtained from 7 different corpora [11-17]. For analyses predicting toddler norms, frequency of the words in input to children up to 30 months was required and so additional frequency counts were obtained from 13 more corpora [18-39]. Transcripts from these 20 corpora covered a wide range of scenarios ranging from structured toy play in a research lab to in-home recordings during typical days. This wide range of scenarios lessened the likelihood that individual objects will be over-represented in their frequency counts due to repeated use by a specific researcher.

Word frequency, semantic density, and sound density were all used in a series of regression analyses predicting the percentage of children that understand or say nouns from the MCDI, at multiple time points.

III. Results

Before beginning, we note that, following Storkel, all significance tests reported here are uncorrected for multiple comparisons. We further note, however, that we only report results that were significant, at an uncorrected level, across a wide and contiguous span of different ages, so that the reported effects are very unlikely to reflect Type-I error.

A. *Testing external validity of the density measure*

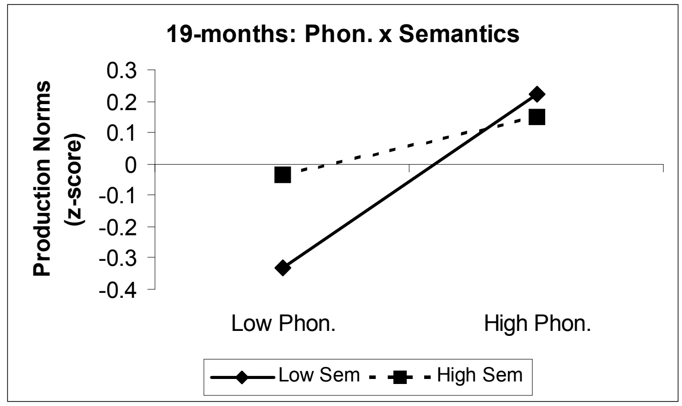
To asses the validity of our new measure of density, we first aimed to replicate Storkel's findings by using word frequency and phonological density to predict the production norms from the MCDI. Just as Storkel did, we found that phonological density was a significant predictor of production norms for ages 14-30 months (all p-values < .05, except for 28 months where p=.066), even after accounting for the effect of word frequency. The standardized coefficients β, for phonological density ranged from .09 - .20 indicating, that as in Storkel's analyses, words from dense phonological neighborhoods were learned earlier than those from sparse neighborhoods.

B. *Analysis of production norms*

The first analysis investigated (1) whether semantic density predicts additional variance in the production-norm data at any age and (2) whether semantic density interacts significantly with phonological density in predicting production-norm data.

To address these questions, we began with simple regressions using word frequency, phonological density, and semantic density to predict proportion of children who know a given word at each age. Frequency and phonological density remained significant positive predictors of production norms even when semantic density was included in the regressions. Frequency was a significant positive predictor for all time points with β ranging from .34-62, while sound density was a significant positive predictor from 14 to16 months for the Infant Says norms and from 16 to 30 months for the Toddler Says norms with β ranging from .09 - .20. Semantic density was a significant predictor only at the 30-month time point with β=.115, indicating that toddlers at this age were more likely to produce words from dense than from sparse semantic neighborhoods. Looking at consecutive ages, the overall r^2 steadily declined with age, from .40-.14. This decline may be explained by an expanding vocabulary and the inability of the MCDI to characterize the vocabulary of older children. At the 30-month time point, only 7% of the words are unknown by more than half of children.

The simple effects suggest that phonological density, but not semantic density, constrains which words children produce earliest. The absence of a simple effect of semantic density does not necessarily mean that semantic density has no influence on lexical acquisition however—it remains possible that these effects interact with phonological and/or frequency effects. To test this possibility, each variable was standardized to a Z-score and the products of variables were computed to examine 2-way and 3-way interactions. We did not



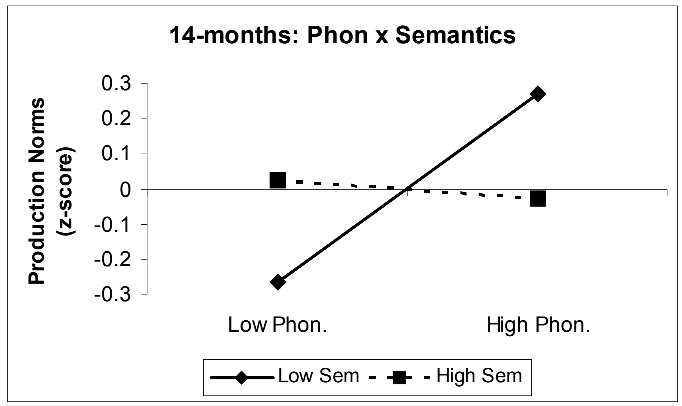


Fig.1:Phonological Density x Semantic Density Interaction in Production at 14– and 19-months.

exhaustively analyze all possible interactions, but focused on two of clear a-priori interest. First, we investigated the interaction between phonological and semantic density; second, we investigated whether this relationship itself varied with word frequency by assessing the 3-way interaction between these factors.

The two-way interaction between phonological density and semantic density was statistically significant in the Infants Says database for all ages between 14 and 16 mos., and in the Toddler Says database between 16 and 19 months. Figure 1 plots the direction of this interaction for the earliest and latest ages in this range; the direction of the interaction was the same throughout this span. Specifically, for words in sparse semantic neighborhoods, phonological density appeared to aid word-learning, whereas words from dense semantic neighborhoods showed a much reduced or absent effect of phonological density. Put differently, phonological density had a facilitating effect for words from sparse semantic neighborhoods, but lesser or no effect on those from dense semantic neighborhoods.

Analysis of the 3-way interaction indicated that this effect was mainly carried by higher-frequency words. This interaction was significant at the oldest age from the Infant Says database and at all ages in the Toddler Says database. Figure 2 plots the 2-way interaction between phonological and semantic density separately for higher and lower-frequency words, for the youngest (16-month) and oldest (30-month) ages in this range. For both 16- and 30-month olds, neither phonological nor semantic density greatly influenced acquisition of lower frequency words; but the factors interacted strongly for higher-frequency words, with phonological density facilitating acquisition of words from sparse but not dense semantic neighborhoods.

In sum: (1) We replicated Storkel's finding that phonological density is associated with earlier word production, using a different metric of phonological density; (2) we found no simple effect of semantic density on acquisition in these norms, but (3) semantic density did seem to modulate the strength and direction of the phonological density effect for higher-frequency words in children ranging from 16-30 months.

C. *Analysis of comprehension norms*

To investigate the degree to which phonological density, semantic density, and word frequency influence the developing ability to comprehend words, we replicated these analyses using lexical norms from the Infant Knows database as our outcome measure.

Unsurprisingly, word frequency was a significant positive predictor at all time points with β ranging from .38-.56. In striking contrast to the previous analysis, semantic density was a significant predictor for all ages between 10 to16 months, with β ranging from .11-.26, whereas phonological density failed to predict lexical norms at any time point.

Analysis of the 2-way interaction between phonological and semantic density revealed a similar pattern to that observed in the production norms: phonological density facilitated comprehension of words from sparse but not dense semantic neighborhoods from 13 to 16 months; the relationship between the predictors was the same at all time points (see Figure 3). In this case, the 2-way interaction was not further modulated by word frequency: the 3-way interaction between these factors was not significant at any time point.

Together, analyses of comprehension and production norms suggest that, even after accounting for the influence of word frequency in speech to children, phonological density is the best predictor of the words children can easily learn to produce; semantic density is the best predictor of the words children can easily learn to understand; and these factors interact in complementary ways, so that phonological effects are most pronounced in words from sparse semantic neighborhoods.

IV. Discussion

In the introduction, we noted an apparent contradiction in previous studies investigating the influence of phonological neighborhood density on lexical acquisition. Specifically, experimental studies of word-learning in the lab have suggested that phonological similarity might hinder word-learning; whereas statistical analyses of large word-production norms have suggested that children are faster to learn words from phonologically dense neighborhoods. We suggested two avenues for reconciling the apparent contradiction. First, we wondered whether the discrepancy might be attributable to an effect of semantic neighborhood density on lexical acquisition. Although much prior research with adults and some research with infants suggests that semantic representations might be

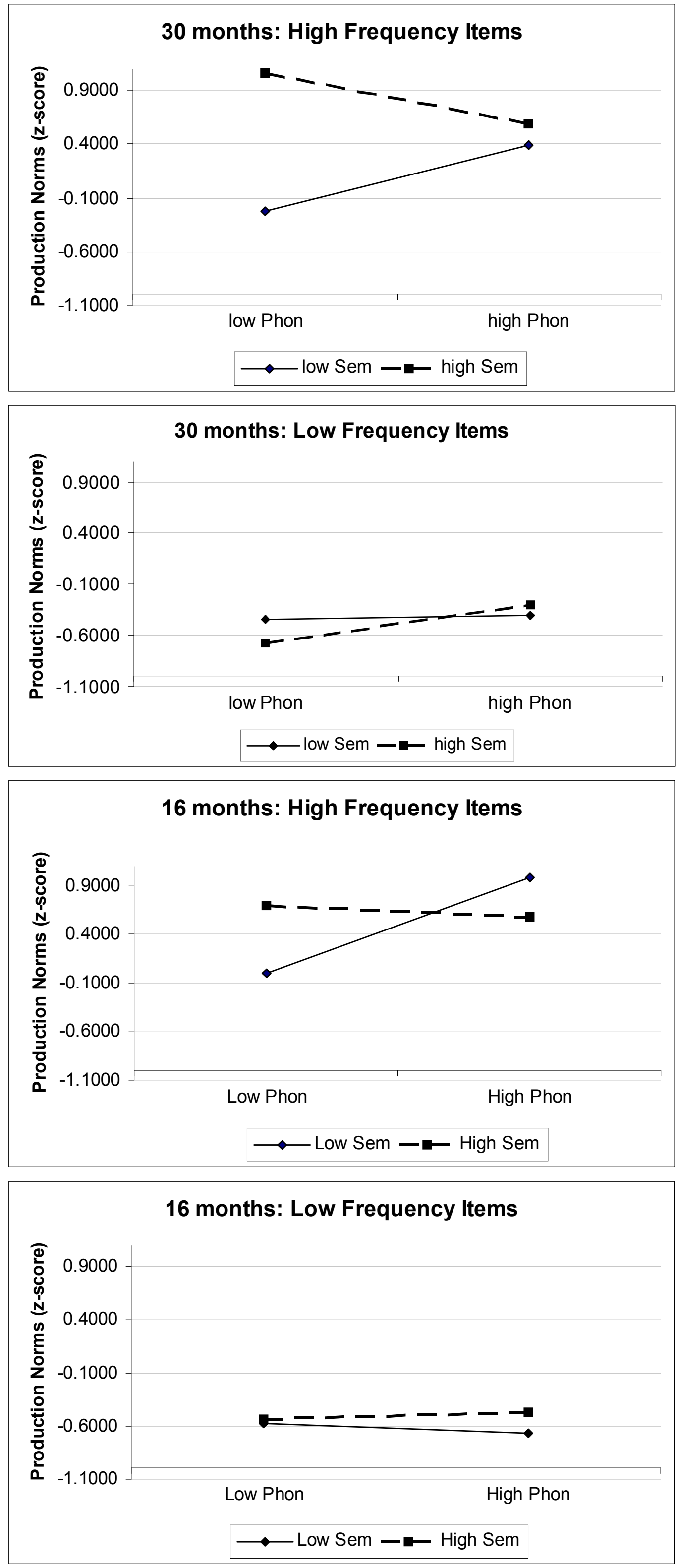


Fig.2:. Phonological Density x Semantic Density Interaction in Production at 14- and 19-months.

an important constraint on the ease with which new words are learned, this is, to our knowledge, the first study to investigate the interaction of semantic and phonological factors in word learning. Second, we noted that the experimental word-

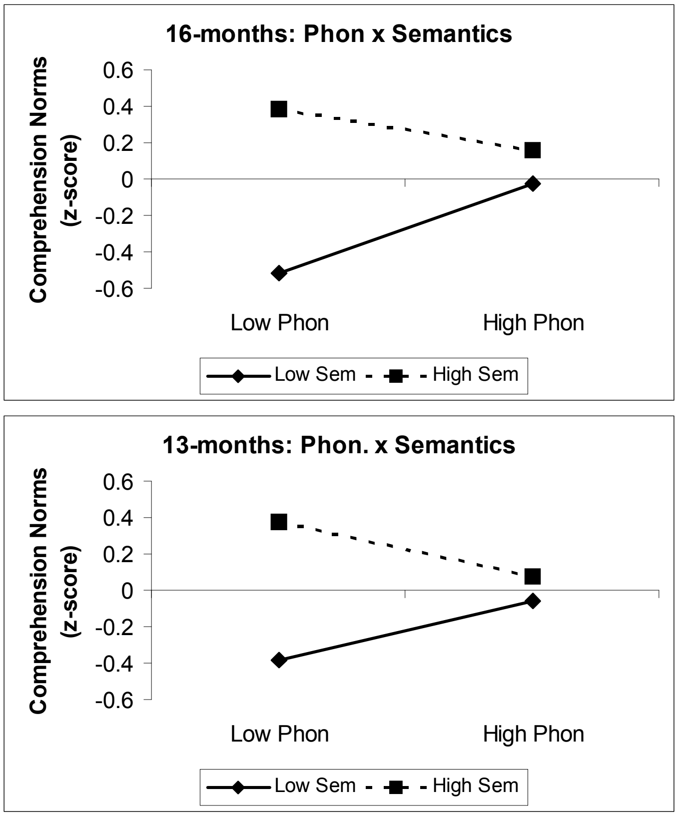


Fig.3:. Phonological Density x Semantic Density Interaction for Comprehension at 13- and 19-months.

learning studies used a receptive measure to assess infant word knowledge whereas the corpus analysis focused solely on production norms. We therefore raised the possibility that phonological information might have different effects on production versus comprehension.

The current results support both of these hypotheses. First, semantic density was shown to modulate the influence of phonological density on the ease with which children learn to produce words. Whereas phonological density boosted acquisition for words from sparse neighborhoods, the same factor had little effect or reversed effect for words from dense semantic neighborhoods. This in turn suggests that the results of experiments like those reported by Werker and by Swingley and Aslin may be strongly influenced by the particular items to which novel labels are mapped and the perceived semantic distinctions amongst them. For instance, if the novel objects are conceived as coming from a dense semantic neighborhood—if they are viewed as kinds of toys, for instance—then one might expect phonological density to hinder acquisition of new words; but if the referents are interpreted as coming from sparse semantic neighborhoods, the reverse finding might hold. In general, the results suggest that, in these experiments, it is important to consider how infants conceive of the referents with which the novel labels are associated.

Second, our analyses revealed very different simple effects: after word frequency, semantic density was the best predictor of lexical comprehension, whereas phonological density was the best predictor of lexical production. Perhaps the apparent contradiction arises because phonological density only facilitates word production, whereas the methods used by Werker and by Swingley and Aslin are clearly more akin to tests of lexical comprehension.

Finally, it is worth noting that, in virtually all analyses, density—whether phonological or semantic—only facilitated lexical acquisition. That is, in production, children learn words from dense phonological neighborhoods prior to words from sparse neighborhoods; and in comprehension, children learn words from dense semantic neighborhoods prior to those from sparse neighborhoods. These findings offer strong constraints on theoretical models of word learning. Specifically, they seem, at least at first blush, to contradict a very basic intuition: it should be harder to learn to distinguish a set of very similar items than a set of very distinct items. Yet this is what the data appear to show. Computational theories of word-learning may offer key insights as to why the basic intuition appears to be wrong.

Appendix A: Sound & Semantic Feature Sets

Sound features: voiced, consonantal, vocalic, sonorant, lateral, continuant, noncontinuant, advanced tongue root, nasal, labial, coronal, anterior, high, distributed, dorsal, radical.

Semantic features: size, weight, strength, speed, temperature, cleanliness, tidiness, brightness, noise, intelligence, goodness, beauty, width, hardness, roughness, height, length, scariness, colorfulness, is black, is blue, is brown, is gold, is green, is grey, is orange, is pink, is purple, is red, is silver, is white, is yellow, is conical, is crooked, is curved, is cylindrical, is flat, is liquid, is rectangular, is round, is solid, is square, is straight, is triangular, has feather, has scales, has fur, is prickly, is sharp, is breakable, made of china, made of cloth, made of leather, made of metal, made of plastic, made of stone, made of wood, climbs, crawls, flies, leaps, runs, swims, breathes, drinks, eats, makes animal noise, singles, talks, has four legs, has beak, has door, has shell, has eyes, has face, has fins, has handle, has leaves, has legs, has paws, has tail, has teeth, has wheels, has whiskers, has wings, is annoying, is comfortable, is fun, is musical, is scary, is strong smelling, is young, is old, is comforting, is lovable, is edible, is delicious.

References

[1] J. F. Werker and C. Fennell, "Listening to Sounds versus Listening to Words: Early Steps in Word Learning," in *Weaving a lexicon.*, D. G. Hall and S. R. Waxman, Eds.: MIT Press, 2004, pp. 79-109.

[2] D. Swingley and R. N. Aslin, "Lexical competition in young children's word learning," *Cognitive Psychology*, vol. 54, pp. 99-132, 2007.

[3] D. Swingley and R. N. Aslin, "Lexical neighborhoods and the word-form representations of 14-month-olds," *Psychological Science*, vol. 13, pp. 480-484, 2002.

[4] D. Swingley and R. N. Aslin, "Spoken word recognition and lexical representation in very young children," *Cognition*, vol. 76, pp. 147-166, 2000.

[5] H. L. Storkel, "Do children acquire dense neighborhods? An investigation of similarity neighborhoods in lexical acquisition," *Applied Psycholinguistics*, pp. 201-221, 2004.

[6] P. S. Dale and L. Fenson, "Lexical development norms for young children.," *Behavior Research Methods, Instruments & Computers*, pp. 125-127, 1996.

[7] K. Graf Estes, "Is there anything wordlike about statistical word segmentation?," in *Workshop on current issues in language acquisition: Artificial languages and statistical learning*. Calgary, Alberta, 2007.

[8] C. Fennell, "Infant attention to phonetic detail in word forms: Knowledge and familiarity effects," in *Department of Psychology*, vol. Doctor of Philosophy. Vancouver, B.C.: The University of British Columbia, 2004, pp. 118.

[9] M. F. Joanisse and M. S. Seidenberg, "Impairments in verb morphology following brain injury: A connectionist model,"

Proceedings of the National Academy of Sciences of the United States of America, pp. 7592-7597, 1999.
[10] S. R. Howell, D. Jankowicz, and S. Becker, "A model of grounded language acquisition: Sensorimotor features improve lexical and grammatical learning," *Journal of Memory and Language*, vol. 53, pp. 258-276, 2005.
[11] B. MacWhinney, *The CHILDES Project: Tools for analyzing talk*, vol. 2: The Databse, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
[12] N. Bernstein, "Acoustic study of mothers' speech to language-learning children: An Analysis of vowel articulatory characteristics.," vol. PhD. Boston: Boston University, 1982.
[13] M. R. Brent and J. M. Siskind, "The Role of Exposure to Isolated Words in Early Vocabulary Development," *Cognition*, vol. 81, pp. 31-44, 2001.
[14] R. P. Higginson, "Fixing-assimilation in Language Acquisition," in *Unpublished doctoral dissertation (Washingtion State University)*: Washington State University, 1985.
[15] A. Ninio, C. Snow, B. Pan, and P. Rollins, "Classifying Communicative Acts in Children's Interactions," *Journal of Communication Disorders*, vol. 27, pp. 157-188, 1994.
[16] J. Sachs, "Talking About the There and Then: The Emergence of Displaced Reference in Parent-child Discourse," in *Children's Language*, vol. 4, K. E. Nelson, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.
[17] L. Bloom, *One Word at a time: The use of single-word utterances before syntax*. The Hague: Mouton, 1973.
[18] A. Warren-Leubecker, "Sex differences in speech to children," Georgia Institute of Technology, 1982.
[19] A. Warren-Leubecker and J. N. Bohannon, "Intonation patterns in child-directed speech: Mother-father speech," *Child Development*, vol. 55, pp. 1379-1385, 1984.
[20] E. Bates, I. Bretherton, and L. Snyder, *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge, MA: Cambridge University Press, 1988.
[21] V. Carlson-Luden, "Causal understanding in the 10-month-old," University of Colorado at Boulder, 1979.
[22] L. Bloom, *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press, 1970.
[23] L. Bloom, L. Hood, and P. Lightbown, "Imitation in language development: If, when and why," *Cognitive Psychology*, vol. 6, pp. 380-420, 1974.
[24] L. Bloom, P. Lightbown, and L. Hood, "Structure and variation in child language," *Monographs of the Society for Research in Child Development*, vol. 40, 1975.
[25] R. Brown, *A first language: The early stages*. Cambridge, MA: Harvard University Press, 1973.
[26] A. Peters, "The role of imitation in the developing syntax of a blind child," *Text*, vol. 7, pp. 289-311, 1987.
[27] B. Wilson and A. M. Peters, "What are you cookin' on a hot?: Movement constraints in the speech of a three-year-old blind child," *Language*, vol. 64, pp. 249-273, 1988.
[28] K. Post, "The language learning environment of laterborns in a rural Florida community," Harvard University, 1992.
[29] K. Post, "Negative evidence," in *Handbook of Research in Language Development Using CHILDES*, J. Sokolov and C. Snow, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994, pp. 132-173.
[30] M. Demetras, K. Post, and C. Snow, "Feedback to first-language learners," *Journal of Child Language*, vol. 13, pp. 275-292, 1986.
[31] V. Valian, "Syntactic subjects in the early speech of American and Italian children," *Cognition*, vol. 40, pp. 21-81, 1991.
[32] E. V. Clark, "Awareness of language: Some evidence from what children say and do.," in *The child's conception of language.*, R. J. A. Sinclair and W. J. M. Levelt, Eds. Berlin: Springer Verlag, 1978.
[33] M. Demetras, "Working parents conversational responses to their two-year-old sons," University of Arizona, 1989.
[34] M. Demetras, "Changes in parents' conversational responses: A function of grammatical development," presented at ASHA, St. Louis, MO, 1989.
[35] M. Demetras, *Working parent's conversational responses to their two-year-old sons*: University of Arizona, 1989.
[36] A. Feldman and L. Menn, "Commentary on 'Filler syllables: What is their status in emerging grammar?' by Ann Peters," *Journal of Child Language*, vol. 28, pp. 269-271, 2001.
[37] A. Feldman and A. F. Healy, "Effect of first language phonological configuration on lexical acquisition in a second language," in *Foreign language learning: psycholinguistic studies on training and retention* A. F. Healy and L. E. Bourne, Eds. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, 1998, pp. 57-76.
[38] P. Suppes, "The semantics of children's language," *American Psychologist*, vol. 29, pp. 103-114, 1974.
[39] L. Van Houten, "Role of maternal input in the acquisition process: The communicative strategies of adolescent and older mothers with their language learning children," presented at Boston University Conference on Language Development, Boston, MA, 1986.