Elsevier Editorial System(tm) for Journal of Memory and Language
Manuscript Draft

Manuscript Number: JML-09-100R3

Title: A Cross-Sectional Comparison of the Effects of Phonotactic Probability and Neighborhood Density on Word Learning by Preschool Children

Article Type: Regular Paper

Keywords: Word Learning; Vocabulary; Phonotactic Probability; Neighborhood Density; Preschool Children.

Corresponding Author: Ms. Jill Hoover, Ph.D.

Corresponding Author's Institution: Indiana University

First Author: Jill Hoover, Ph.D.

Order of Authors: Jill Hoover, Ph.D.; Holly L Storkel; Tiffany P Hogan

Abstract: Two experiments examined the effects of phonotactic probability and neighborhood density on word learning by 3-, 4-, and 5-year-old children. Nonwords orthogonally varying in probability and density were taught with learning and retention measured via picture naming. Experiment 1 used a within-story probability/across-story density exposure context. Experiment 2 used an across-story probability/within-story density exposure context. Results showed that probability and density interacted to create optimal learning conditions. Specifically, rare/sparse sound sequences appeared to facilitate triggering of word learning. In contrast, the optimal convergence for lexical configuration and engagement was dependent on exposure context. In particular, common sound sequences and dense neighborhoods were optimal when density was manipulated across stories, whereas rare sound sequences and sparse neighborhoods were optimal when density was manipulated within a story. Taken together, children's phonological and lexical representations were hypothesized to be interdependent on one another resulting in a convergence of form characteristics for optimal word learning.

Running Head: PRESCHOOL WORD LEARNING

A Cross-Sectional Comparison of the Effects of Phonotactic Probability and Neighborhood Density on Word Learning by Preschool Children

Jill R. Hoover

Indiana University

Holly L. Storkel

University of Kansas

Tiffany P. Hogan

University of Nebraska - Lincoln

Under Review, *Journal of Memory and Language*

Contact author: Holly L. Storkel, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas, 3001 Dole Human Development Center, 1000 Sunnyside Ave., Lawrence, KS 66045-7555. Email: hstorkel@ku.edu, Phone: (785) 864-0497, Fax: (785) 864-3974.

Abstract

Two experiments examined the effects of phonotactic probability and neighborhood density on word learning by 3-, 4-, and 5-year-old children. Nonwords orthogonally varying in probability and density were taught with learning and retention measured via picture naming. Experiment 1 used a within-story probability/across-story density exposure context. Experiment 2 used an across-story probability/within-story density exposure context. Results showed that probability and density interacted to create optimal learning conditions. Specifically, rare/sparse sound sequences appeared to facilitate triggering of word learning. In contrast, the optimal convergence for lexical configuration and engagement was dependent on exposure context. In particular, common sound sequences and dense neighborhoods were optimal when density was manipulated across stories, whereas rare sound sequences and sparse neighborhoods were optimal when density was manipulated within a story. Taken together, children's phonological and lexical representations were hypothesized to be interdependent on one another resulting in a convergence of form characteristics for optimal word learning.

Keywords: Word Learning; Vocabulary; Phonotactic Probability; Neighborhood Density; Preschool Children.

Phonological and lexical representations are two of the types of representations that have been hypothesized to play a role in word learning (e.g., Gupta & MacWhinney, 1997). P*honological* representations refer to the individual sounds in a word form (e.g., /d/ /ɑ/ /g/). L*exical* representations refer to a word form in its entirety (e.g., /dɑg/). When a novel word form is encountered phonological representations will be activated, assuming the sounds are part of the learner's phonology. Lexical representations of similar sounding words will also be activated. The novel word form must be recognized as novel, rather than known, so that learning will be triggered. Once learning is triggered, lexical configuration, or the creation of a *new* lexical representation, occurs (Leach & Samuel, 2007). Repeated exposure to the new word along with subsequent activation of lexical representations will also lead to lexical engagement, which involves the integration of the new representation with existing lexical and phonological representations (i.e., creation of links between similar representations) and takes place over time (Leach & Samuel, 2007). This model explains the general components of word learning for any learner. However, it is likely that the characteristics that influence each aspect of the word learning model may change across development. Thus, the developmental version of the model leads to a focus on the characteristics that are used by learners at different points in development as well as the relative weighting of these characteristics at different points in development (Hirsh-Pasek, Golinkoff, & Hollich, 2000).

Two form characteristics that have received recent attention for their role in word learning are phonotactic probability and neighborhood density. *Phonotactic probability* refers to the likelihood of occurrence of individual sounds and sound sequences in a language and is a characteristic of phonological representations (Vitevitch & Luce, 1999). Sounds and sound

combinations such as those in the word *cat* are highly likely to occur and are consequently referred to as *common*. On the other hand, sounds and sound combinations such as those in the word *cheese* are less likely to occur and are consequently referred to as *rare*. *Neighborhood density* is the number of words that are phonologically similar to a given word based on a one sound substitution, addition, or deletion and is a characteristic of lexical representations (Vitevitch & Luce, 1999). Words such as *coat* have many similar sounding words and consequently reside in dense neighborhoods. On the other hand, words such as *these* have few similar sounding words and reside in sparse neighborhoods.

Although phonotactic probability and neighborhood density are two distinct characteristics, they are highly related (Storkel, 2004b; Vitevitch, Luce, Pisoni, & Auer, 1999). Specifically, there is a significant positive correlation between phonotactic probability and neighborhood density in English in that words like *cat* that are composed of common sound sequences also tend to reside in dense neighborhoods (Storkel, 2004b). In complement, words like *cheese* that are composed of rare sound sequences tend to reside in sparse neighborhoods. For this reason, stimuli orthogonally varying in phonotactic probability and neighborhood density must be presented in learning tasks to completely understand the individual, and combined, effects of phonological and lexical representations on word learning.

One study of adult word learning that orthogonally varied phonotactic probability and neighborhood density suggested that both characteristics were used to learn new words, but that each influenced a different component of the learning process. Specifically, results of Storkel, Armbruster and Hogan (2006) showed significant main effects of both phonotactic probability and neighborhood density and no interaction between the two characteristics. Specifically, adults learned rare sound sequences more readily than common, and this effect was apparent early in

word learning. In contrast, dense words were learned more readily than sparse words, and this effect was observed later in word learning.

Converging results were obtained using linear regression to determine when infants learn words (Storkel, 2009). Specifically, infants learned real words composed of rare sound sequences at an earlier age than real words composed of common sound sequences, and this effect did not interact with age. In addition, infants learned real words from dense neighborhoods at an earlier age than real words from sparse neighborhoods, and this effect did interact with age. Specifically, the influence of neighborhood density on word learning decreased as age increased. It is important to note that the potential interaction between phonotactic probability and neighborhood density was not examined in this infant study nor was it possible to examine whether the effect of each variable occurred early or late in the word learning process because only the outcome of word learning (i.e., known words) was examined rather than the dynamic process of word learning itself.

Taken together, these results led to the hypothesis that phonological representations, as indexed by phonotactic probability, were critical to triggering the word learning process (i.e., identifying a word as novel so as to initiate learning). Since rare sound sequences are not encountered in the ambient language as frequently as common sound sequences, they will immediately be identified as novel with learning triggered sooner than for common sound sequences. The onset of learning may be delayed for common sound sequences because they are deceptively similar to other known sequences in the language. Furthermore, it was hypothesized that the role of phonotactic probability in triggering word learning was established early in development and potentially remained stable throughout development. In terms of the dense advantage, it was hypothesized that lexical representations, as indexed by neighborhood density,

played more of a role in lexical configuration and/or engagement (Leach & Samuel, 2007; Storkel, in press). Dense words place fewer demands on working memory ability because they are easier to maintain compared to sparse words. Thus the connections inherent to dense words support the creation of an accurate and detailed lexical representation during configuration (e.g., Roodenrys & Hinton, 2002; Thomson, Richardson, & Goswami, 2005; Thorn & Frankish, 2005). In terms of engagement, novel words from dense neighborhoods are connected with many more existing lexical representations than novel words from sparse neighborhoods, potentially anchoring and strengthening the lexical representation of the new word. In terms of developmental effects, it was hypothesized that the role of neighborhood density in configuration and engagement potentially changed over time.

One gap in the application of this model to development is that interactions between phonotactic probability and neighborhood density across ages have not been investigated. It is possible that the role of phonotactic probability and neighborhood density in word learning is not strictly circumscribed to one component of word learning, as the evidence suggests for adults. That is, children may require a convergence of multiple form characteristics to support efficient triggering, configuration, and engagement. No empirical study, like Storkel et al. (2006), has been carried out with children to provide the necessary evidence to address this issue. Rather, all experimental word learning studies to date have examined phonotactic probability when correlated with neighborhood density, and have tended to study a narrow age window (i.e., infants) or have collapsed across ages (i.e., preschool children). These studies provide some initial insights into the role of phonotactic probability and neighborhood density in word learning by children, but also raise additional questions.

In terms of infant word learning studies, the effect of correlated phonotactic probability and neighborhood density has been inconsistent. In one such study, the phonotactic probability and neighborhood density of the to-be-learned words was established in a pre-exposure condition by presenting either many words that were similar to the to-be-learned words (i.e., common/dense) or few words that were similar to the to-be-learned words (i.e., rare/sparse). With only a brief pre-exposure to the similar words, 17-month old infants learned a new word composed of common/dense, rather than rare/sparse, sound sequences (Hollich, Jusczyk, & Luce, 2002). However, when pre-exposure was longer, entailing more repetitions of the similar words, infants learned a rare/sparse word, but not a common/dense word. In a similar vein, a more recent study using eye-tracking suggested that 18-month-old infants failed to learn neighbors of known words, although they were able to learn words that are dissimilar to all known words (Aslin & Swingley, 2007). Another study with slightly older infants (i.e., 20- and 24-month-olds) showed that learning was better for a common/dense, rather than a rare/sparse word for children when neighborhood size of the to-be-learned word was determined by considering existing words in the child's vocabulary (Newman, Samuelson, & Gupta, 2008). The findings across these infant studies suggest that the effect of phonotactic probability and neighborhood density on word learning by young children may vary by context (i.e., experimental paradigm) and possibly by age.

Turning to studies with preschool children, common/dense novel words were learned more readily than rare/sparse novel words (Storkel, 2001, 2003; Storkel & Maekawa, 2005; Storkel & Rogers, 2000). Moreover, this finding showed that vocabulary development inconsistently influenced word learning. Specifically, one study showed that the size of the common/dense advantage increased as vocabulary increased (Storkel, 2001), whereas another

failed to detect this same relationship (Storkel, 2003). Taken together, the results from preschool children are consistent with neighborhood density findings from the past study of adult word learning (i.e., dense learned better than sparse), suggesting that neighborhood density may play a similar role in word learning by young children. However, because only correlated phonotactic probability and neighborhood density were examined, it was unclear whether both characteristics individually or collectively influenced word learning by preschool children, and whether each characteristic had a unique influence on a particular component of word learning (i.e., triggering, configuration, engagement), as in adults. Moreover, the influence of development during the preschool period on the effects of phonotactic probability and neighborhood density is unclear, with conflicting results across studies.

The inconsistent effects of form characteristics on word learning by infants suggest that the manner in which words varying in phonotactic probability and neighborhood density are presented to the learner might differentially affect learning. For example, in infant word learning the amount of exposure to similar sounding novel words leads to differing effects of phonotactic probability and neighborhood density, as previously described (Hollich et al., 2002). This suggests a need to examine the influence of exposure paradigms (i.e., context) on the effects of phonotactic probability and neighborhood density on word learning. Past studies of word learning by preschool children have tended to use the same context (Storkel, 2001, 2003, 2004a; Storkel & Maekawa, 2005). Specifically, novel words that vary in correlated phonotactic probability/neighborhood density (i.e., rare/sparse or common/dense) are paired with referents from the same semantic category (e.g., toys) and then are presented to children simultaneously during training. That is, a child might hear a rare/sparse novel word paired with one toy during training and then immediately hear a common/dense novel word paired with another toy. In this

scenario, it is likely easier to hold the common/dense sound sequence in working memory compared to the rare/sparse sound sequence, leading to creation of a more accurate and detailed representation for the common/dense sound sequence. However, it is unclear whether this same pattern would be obtained if the rare/sparse novel word was not presented in direct opposition to the common/dense novel word. Thus, the generality of phonotactic probability and neighborhood density effects across different contexts needs to be examined.

Purpose of the Current Studies

The goal of the current studies was to disentangle the effects of phonotactic probability and neighborhood density on word learning by preschool children varying in age. Accordingly, stimuli were constructed to orthogonally vary phonotactic probability and neighborhood density so that the individual and interactive effects of these two characteristics on word learning by preschool children could be examined. In addition, word learning was measured at two time points: 1) immediately following exposure and 2) one-week after exposure to determine whether similar effects of phonotactic probability and neighborhood density were observed across time, which could reveal the specific components of word learning (i.e., triggering, configuration, or engagement) that are influenced by each form characteristic. Developmental issues also were examined by comparing the effects of phonotactic probability and neighborhood density across 3- versus 4- versus 5-year-old children and to past studies of adult word learning (i.e., Storkel et al., 2006) to determine whether the effect of each characteristic is stable or variable across the lifespan. Lastly, contextual influences were examined across two studies to determine whether the effect of each characteristic was stable or variable across contexts to inform whether learning of certain words is more or less difficult depending on the learning context. Specifically, Experiment 1 directly paired nonwords varying in phonotactic probability but matched in

neighborhood density within the same story, whereas the neighborhood density manipulation occurred across stories. In complement, Experiment 2 directly paired nonwords varying in neighborhood density but matched in phonotactic probability within the same story, whereas the phonotactic probability manipulation occurred across stories.

If word learning by preschool children is similar to adults, a robust effect of phonotactic probability and neighborhood density without an interaction is predicted. In this case, children would show higher accuracy for rare sound sequences than for common, regardless of neighborhood density, and would show higher accuracy for dense neighborhoods than for sparse, regardless of phonotactic probability. Moreover, phonotactic probability would tend to show effects at the immediate time point, indicating a role in triggering, and neighborhood density would tend to show effects at the 1-week time point, indicating a role in configuration and engagement (Leach & Samuel, 2007; Storkel, in press). In contrast, if children require a convergence of characteristics for efficient word learning, then an interaction of phonotactic probability and neighborhood density would be observed. Specifically, at the immediate time point, early in word learning during triggering, rare/sparse sound sequences may be higher in accuracy than other sound sequences because both phonotactic probability and neighborhood density converge to indicate that the sound sequence is novel. At the 1-week time point, later in word learning during configuration and engagement, common/dense sound sequences may be higher in accuracy than other sound sequences because both phonotactic probability and neighborhood density converge to support maintenance of the novel sound sequence in working memory (Gathercole, Frankish, Pickering, & Peaker, 1999; Roodenrys & Hinton, 2002; Thomson et al., 2005; Thorn & Frankish, 2005). Either prediction is in line with the main effects

of Storkel (2009). The crucial issue is the absence or presence of an interaction between phonotactic probability and neighborhood density.

Predictions concerning developmental and context effects are less clear based on past literature. Developmental differences have been inconsistent. Thus, a lack of interaction between phonotactic probability/neighborhood density and age as well as a significant interaction between phonotactic probability/neighborhood density and age are equally likely based on past work. In terms of context, infant studies have shown variability in the effect of phonotactic probability/neighborhood density across different word learning paradigms but preschool studies have tended to use a single paradigm without varying context. Therefore, it is unknown whether the effect of phonotactic probability or neighborhood density will be robust to variations in context.

Experiment 1

Experiment 1 contrasted nonwords against one another within a story so that one nonword was composed of a common sound sequence and the other was composed of a rare sound sequence, while neighborhood density was held constant. This pairing also resulted in the contrasting of nonwords against one another across stories so that nonwords within a semantic category in one story were dense while the nonwords in the same semantic category in the other story were sparse. In other words, Experiment 1 allowed for a comparison of common versus rare sound sequences *within* a story (i.e., phonotactic probability) and dense versus sparse nonwords *across* two stories (i.e., neighborhood density).

Methodology

Participants. Three groups of typically developing children were recruited for this study: twenty-three 3-year-olds (14 boys and 9 girls*)*, thirty-three 4-year-olds (20 boys and 13 girls*)*,

and twenty-three 5-year-olds (14 boys and 9 girls). The mean and range of ages within each group are shown in Table 1.

All participants in this and the following experiment were recruited from Lawrence, Topeka, and greater Kansas City, Kansas. All children in this and the following experiment were monolingual native English speakers. Typical development for both experiments was verified through (1) performance within the normal limits on standardized measures of either receptive and/or expressive vocabulary development (Brownell, 2000a, 2000b); (2) performance within the normal limits on a standardized measure of phonology (Goldman & Fristoe, 2000); (3) a normal hearing screening (ASHA, 1997). The mean, standard deviation and range of scores for each age group in both experiments are shown in Table 1. Production of sounds used in the nonword stimuli was further assessed by examining words on the *Goldman Fristoe Test of Articulation-Second Edition* (*GFTA-2*; Goldman & Fristoe, 2000) and on a supplemental picture naming probe constructed specifically for this study. The supplemental picture naming probe assessed each sound in the target word position (i.e., word-initial or word-final position) in familiar real words that were not presented on the *GFTA-2*. Correct production of the target sounds was required to guard against misarticulation of the nonword stimuli during word learning.

Materials. Nonword stimuli were comprised of 16 consonant-vowel-consonant (CVC) nonwords composed of early acquired consonants (i.e., glides, anterior nasals, and anterior stops). Stimuli differed on two independent variables: phonotactic probability and neighborhood density. The same procedures for selecting nonwords in Storkel et al. (2006) were used in this study. Phonotactic probability and neighborhood density were initially computed using the Hoosier Mental Lexicon (HML), a 20,000-word computerized dictionary containing phonemic transcriptions, word familiarity ratings (Nusbaum, Pisoni, & Davis, 1984) and word frequency

(Kucera & Francis, 1967). After the stimuli were selected, an on-line child calculator became available (http://www.bncdnet.ku.edu/cml/info_ccc.vi that used these same algorithms to calculate phonotactic probability and neighborhood density using kindergarten and first grade child corpora (Kolson, 1960; Moe, Hopkins, & Rush, 1982). Table 2 presents the phonotactic probability and neighborhood density values obtained from both the HML and the child calculator. In general, the child values resulted in a similar classification of the stimuli.

Phonotactic probability. Two measures of phonotactic probability were computed following previously documented procedures (Storkel, 2004b; Vitevitch & Luce, 1998): positional segment frequency and biphone frequency. Positional segment frequency is the likelihood of occurrence of a single sound in a given word position (e.g., the likelihood that /k/ in the word /kæt/ occurs in the first word position). The positional segment frequency of each nonword was computed by summing the positional segment frequencies of each individual sound in the nonword. To compute the positional segment frequency for one sound, the sum of the log frequency of all words in a corpus (i.e., HML or child corpus) containing the target sound (e.g., /k/) in the target word position (e.g., first position) was divided by the sum of the log frequency of all words in the corpus containing any sound (e.g., sounds other than /k/) in the target word position (e.g., first position). Thus, positional segment frequency is a measure of relative frequency.

Biphone frequency is the likelihood of occurrence of two adjacent sounds (e.g., the likelihood that the sequence /kæ/ in the word /kæt/ occurs in the first position). The biphone frequency of each nonword was computed by summing the biphone frequencies of each pair of sounds in the nonword. To compute the biphone frequency for one pair of sounds, the sum of the log frequencies for all words in a corpus (i.e., HML or child corpus) containing the target pair in

the target word position (e.g., /kæ/ in the first position, as in “cab”, “cap” “can”) was divided by the sum of the log frequencies of the words in the corpus containing any sound in the target word position. Thus, biphone frequency is a measure of relative frequency.

Positional segment frequency and biphone frequency were computed for all legal CVC nonword sequences according to American English. Following the procedures of Storkel (2004), a median split, based on the available pool of possible CVC stimuli, was used to categorize each CVC as having either common or rare sound sequences. Nonwords with both a positional segment frequency and biphone frequency value above the median of all possible CVC stimuli in the available stimuli pool were coded as common whereas patterns with both a positional segment frequency and biphone frequency value below the median of all possible CVC stimuli in the available stimuli pool were coded as rare.

Neighborhood density. Neighborhood density was computed by counting the number of words in a corpus (i.e., HML or child corpus) that differed from a given CVC nonword by a one sound substitution, addition, or deletion. A median split based on the available pool of CVC nonwords was used to categorize each CVC as either dense or sparse. CVC nonwords that had more neighbors than the median value of the CVC stimuli pool were coded as residing in a dense neighborhood whereas CVC nonwords with fewer neighbors than the median value of the CVC stimuli pool were coded as residing in a sparse neighborhood.

The current study selected the previously described algorithms for computing phonotactic probability and neighborhood density to afford comparisons with past word learning studies that have used these same algorithms (Storkel, 2001, 2003; Storkel, 2009; Storkel et al., 2006; Storkel & Maekawa, 2005). It is important to note that these algorithms make certain assumptions that have yet to be conclusively verified through empirical study. For example, both

algorithms are sensitive to word position but insensitive to syllable structure, in spite of evidence from psycholinguistic studies suggesting that humans are sensitive to the internal structure of words (e.g., Treiman, Fowler, Gross, Berch, & Weatherson, 1995; Ziegler & Goswami, 2005). Few studies have compared different algorithms for computing phonotactic probability and neighborhood density (but see Bailey & Hahn, 2001; Frisch, Large, & Pisoni, 2000) to determine which algorithm is the best predictor of human performance. In the absence of such evidence, the current study opted for continuity with past research on word learning, but this does not entail unequivocal endorsement of the algorithms used.

The 16 nonwords shown in Table 2 were equally divided into four conditions based on their phonotactic probability and neighborhood density classifications: (a) common/dense, (b) common/sparse, (c) rare/dense, and (d) rare/sparse. Across the conditions where phonotactic probability varied but neighborhood density was balanced, the number of neighbors was similar. Similarly, across the conditions where phonotactic probability was balanced but neighborhood density varied, the average positional segment frequency and biphone frequency values were similar. It should be noted that fine grain acoustic properties of our nonword stimuli were not systematically matched across the phonotactic probability and neighborhood density conditions. Research examining acoustic factors has in fact noted that real words with dense neighborhoods tend to be produced with a more expanded vowel space, which might inherently enhance stimuli intelligibility and consequently enhance word learning ability (e.g., Munson & Solomon, 2004). This issue warrants attention in future studies.

Novel object referents. Following the procedures of Storkel et al. (2006), each nonword selected for this study was arbitrarily paired with a picture of a novel object referent. The same novel object referents that were used in Storkel et al. (2006) were used in this study. Novel

object referents did not have a corresponding label in English and thus were not identifiable by children enrolled in the study. The novel objects were either created or adapted from children's stories (DeBrunhoff, 1981; Geisel & Geisel, 1954, 1958; Mayer, 1992). The 16 novel object referents came from four semantic categories: toys, horns, candy machines, and pets. The four semantic categories were matched across the four phonotactic probability/neighborhood density conditions so that each condition contained one nonword-novel object referent pair from each semantic category.

Stories. The same two stories used in Storkel et al. (2006) were also used in this study however the number of exposures to each nonword was increased in the current study to guard against floor effects in children. The 16 nonword-novel object referent pairs were divided into two sets of eight with the semantic categories balanced across the sets. Therefore, within each set of nonwords, two toys, two pets, two horns, and two candy machines were presented. Each set of eight was presented within a separate story on separate days with approximately 1 week in between each presentation. In both stories, two nonwords were assigned to each semantic category so that each category held neighborhood density constant, while varying phonotactic probability (e.g., Toy 1 paired with a common/sparse nonword vs. Toy 2 paired with a rare/sparse nonword). In other words, within a story each semantic category (e.g., toys) contrasted a common/sparse nonword with a rare/sparse nonword or a common/dense nonword with a rare/dense nonword. Across stories, this pairing contrasted a semantic category paired with sparse nonwords against the same semantic category paired with dense nonwords and vise versa. This assignment allows for two comparisons: 1) common versus rare sound sequences within a semantic category and *within* a story and 2) dense versus sparse nonwords within a semantic category, but *across* a story. This procedure for pairing nonwords with novel object

referents within and across stories is the critical difference between Experiment 1 and Experiment 2. The order of presentation of the stories was counterbalanced across participants.

Stories consisted of three distinct episodes where six visual scenes were presented per story episode (introduction, four intermediate scenes, and a conclusion). The visual scenes presented were adapted from children's books (Mayer, 1993). All novel object referents were embedded into the visual scenes. All visual scenes were presented with a corresponding audio narrative script. Each story introduced children to two main characters and one main event likely to be familiar to young children (e.g., boy and girl character participating in show-and-tell). Following the introduction, four intermediate scenes with corresponding audio narrative script provided the exposure to the nonword-novel object referent pairs. The two nonword-novel object referent pairs from a given semantic category were presented simultaneously within one scene (e.g. toy 1 followed by toy 2 in scene 1). Each intermediate scene featured the main characters interacting with each nonword-novel object referent pair. Nonwords were embedded within a sentence and incorporated into the audio narrative script. Following the four intermediate scenes, a conclusion to the main activity was presented. Characters (e.g., boy and girl) remained the same across each of the three story episodes; however the main activity changed across episodes (e.g., going to the park with objects, competing against each other using objects, playing hide-and-seek with objects, deciding what to bring for show-and-tell, participating in show-and-tell, finding lost objects after show-and-tell). Each nonword was presented four times within a story episode. Following each story episode the nonwords were reviewed one by one in an elicited production task (e.g., "Look, it's a *woʊn*, Say *woʊn*, Remember it's a *woʊn"*). Therefore, after episode 1, children were exposed to each nonword eight times, 16 times after episode 2, and 24 times after episode 3. Following a 1-week delay from the initial 24 exposures, each of the

nonword-novel object referent pairs was reviewed one final time in the elicited production task resulting in 28 total exposures.

All visual scenes were digitized and edited. The audio narrative script was recorded in a soundproof booth, digitized, and edited using the Computerized Speech Lab software. The speaking rate used in the recording of the stories, as measured in syllables per second, was similar across the phonotactic probability/neighborhood density conditions F (2, 120) < 1.0, p = .984. Nonword stimuli and audio quality were verified via the transcription of each stimulus presented in the story and elicited production scripts under the same listening conditions as the participants completed by two blind judges.

Measures of learning. Learning was measured using the same picture naming task used in Storkel et al. (2006). In this task, children were shown a picture of the novel object referent and were asked to produce the corresponding nonword. This task was administered five times per story: Prior to the story to obtain a baseline measure, immediately following each of the three story episodes, and 1 week following the story to obtain a post measure of learning. Responses were phonetically transcribed and scored. Responses were scored as correct if the child's production included at least two of the three phonemes in the correct word position ignoring phoneme additions and deletions (e.g., *waʊni* for *woʊn* would be scored as correct). Proportion correct for each phonotactic probability/neighborhood density condition at 24 (i.e., immediately following the end of the story) and 28 exposures (i.e., 1-week after the story) served as the dependent variable for all analyses. Other test points were not analyzed due to potential floor effects that could vary by age.

Procedure

Each child was seated in front of a laptop computer connected to desktop speakers. Children's responses were recorded using a head-mounted microphone, a digital tape recorder, and a video recorder. Auditory and visual stimuli were presented and controlled by the laptop computer using DirectRTv.2006 software (Jarvis, 2002).

The study required four, 45-minute sessions. The first session was used to screen the child's articulation using the *GFTA-2* (Goldman & Fristoe, 2000) and the supplemental picture naming probe, and to assess hearing to determine study eligibility.

The second session began with the administration of the picture naming task to obtain the baseline measure. Children were told that they would see objects that they had never seen before and were instructed to guess the name of each object. All responses were phonetically transcribed and audio recorded. Next, the first episode of the story was presented over the desktop speakers. The introductory and conclusion scenes were always presented at the beginning and end of each story episode. The order of presentation of the four intermediate scenes providing exposure to the nonword-novel object pairs was randomized as determined by the Direct RT software. The random order of the scenes did not interfere with the cohesiveness of the story because each scene was related to an overarching event (e.g., show-and-tell) and made no reference to any of the other scenes presented in the story. Following each story episode, additional exposures to each nonword-novel object pair were provided via elicited production. Following the elicited production exposures, learning of the nonword-novel object referent pairs was measured via the picture naming task. Children were instructed to try to recall the names of the objects as they were presented in the story. The same procedure was followed for the second and third episodes of the story.

The third session began by administering the elicited production and picture naming tasks from the first story. The participants were instructed to try to recall the names of the objects from the story that they had heard last time. After completing the elicited production and picture naming task from the first story, children were shown the second set of novel objects that would be presented in the new story (i.e., picture naming baseline for the second story). The procedures for the second story mirrored those of the first story.

One week later, the elicited production and picture naming task for the second story were administered during the fourth session. Following these tasks, the receptive and expressive vocabulary tests were administered.

Reliability

Transcription reliability for consonants was computed for 22% of the sample. Transcription reliability was calculated for real word productions made on the *GFTA-2* and for nonword productions during the word learning protocol. Inter-judge transcription reliability for real words was 96% (*SD = 2.7%*, range = 91% to 100%). Inter-judge transcription reliability for nonwords was 96% (*SD* = 4%, range = 84% to 100%).

Scoring reliability was computed for 25% of the sample. Reliability was calculated for scoring the child's production against the target nonword (word score) and for classifying the child's response as either correct or incorrect (correct score). Inter-judge scoring reliability for the nonword score was 98% (*SD* = 2.3%, range = 92% - 100%). Inter-judge scoring reliability for the correct score was 98% (*SD* = 2.7%, range = 91% - 100%).

Data collection procedural reliability was computed for 27% of the sample. Procedural reliability was computed to ensure that the same set of procedures was followed across all data collectors. Inter-judge procedural reliability was 96% (*SD* = 4%, range = 87% - 100%).

Statistical Analyses

The dependent variable was the proportion of nonwords correct in the picture naming task for each phonotactic probability/neighborhood density condition (i.e., common/dense, common/sparse, rare/dense, and rare/sparse) at two time points: (1) immediately after exposure and (2) 1-week after exposure. The main analysis used for this study was a 2 (within story phonotactic probability) x 2 (across story neighborhood density) x 2 (time) x 3 (age) repeated measures analysis of variance (ANOVA) with planned age comparisons.

To interpret significant interactions (i.e., $p < .05$), the effect of the one variable involved in the interaction (e.g., phonotactic probability) was explored for each level of the second variable involved in the interaction (e.g., sparse versus dense neighborhood density) using multiple ANOVAs. In complement, the effect of the second variable involved in the interaction (e.g., neighborhood density) was explored for each level of the first variable involved in the interaction (e.g., rare versus common phonotactic probability) using multiple ANOVAs. Three-way and four-way interactions were explored in a similar manner. For example, to interpret significant three-way interactions (e.g., phonotactic probability x time x age), the effect of two variables involved in the interaction (e.g., phonotactic probability and age) would be examined at each level of the third variable (e.g., time: immediate vs. 1-week retention) and any significant two-way interactions within these separate ANOVAs (e.g., phonotactic probability x age) would be unpacked following the previously described procedures for two-way interactions. As recommended by Levin, Serlin, and Seaman (1994), alpha was held at .05 for each follow-up ANOVA involving interactions with at least one variable involving only two levels. For any interaction involving age, which has three levels, *t* test comparisons were used to examine effects

of neighborhood density or phonotactic probability at each age with alpha held at the .05 level for determining significance (Levin, Serlin, & Seaman, 1994).

In general, all results are reported for the main ANOVA. The presentation of results of the follow-up analyses focuses solely on significant main effects and interactions that are critical for the research questions (i.e., those main effects and interactions involving phonotactic probability and neighborhood density). Thus, significant main effects of time and significant interactions between time and age will be reported, but will not be explicitly discussed because they are not directly relevant to the research questions. Non-significant effects generally are not reported for follow-up analyses, all Fs < 2.74, all ps > .05, all η_p^2 s < .08.

Results

Table 3 presents the raw accuracy data (i.e., means, standard deviations, and 95% confidence intervals) for all four experimental conditions (i.e., common/dense, common/sparse, rare/dense, rare/sparse) at each time (i.e., immediate and retention) and for each age (i.e., 3-, 4-, and 5- years).

The main 2 (phonotactic probability) x 2 (neighborhood density) x 2 (time) x 3 (age) ANOVA showed a significant main effect of time, $F(1, 76) = 32.16$, $p < .001$, $\eta_p^2 = .297$, with higher accuracy at the 1-week retention test (M = .20, SD = .21, SEM = .02) compared to the immediate test (M = .13, SD = .17, SEM = .02). The main effect of phonotactic probability also was significant, $F(1,76) = 6.31$, $p < .05$, $\eta_p^2 = .077$, with higher accuracy for rare (M = .18, SD = .20, SEM = .02) than common sound sequences (M = .15, SD = .17, SEM = .02). Lastly, there was a main effect of neighborhood density $F(1, 76) = 4.69$, $p < .05$, $\eta_p^2 = .058$, with higher accuracy for dense (M = .18, SD = .16, SEM = .02) than sparse neighborhoods (M = .15, SD = .19, SEM = .02). The main effect of age was not significant, with 3-year-olds (M = .13, SD =.19,

SEM = .03) performing similarly to 4-year-olds (*M* = .17, *SD* = .19, *SEM* = .02) who both performed similarly to 5-year-olds (*M* = .20, *SD* = .20, *SEM* = .03), $F(2, 76) = 2.37$, $p = .101$, η_p^2 .06. The significant main effects were qualified only by a significant interaction between phonotactic probability and neighborhood density $F(1, 76) = 5.61$, $p < .05$, $\eta_p^2 = .07$. The interactions involving age with phonotactic probability, neighborhood density, or time were not significant, all *F*s < 2.81, all *p*s > .180, all η_p^2s < .07. Likewise, none of the interactions involving time with age, phonotactic probability, or neighborhood density were significant, all *F* s < 2.9, all *p*s > .07, all η_p^2s < .07. Follow up ANOVAs were conducted as previously described to determine the pattern of effects of phonotactic probability for dense and sparse nonwords and to determine the pattern of effects of neighborhood density for common and rare sound sequences.

The first follow up ANOVA to examine the effect of phonotactic probability for sparse nonwords showed a significant main effect time $F(1, 76) = 12.84$, $p < .01$, $\eta_p^2 = .145$, which mirrored that described for the main ANOVA (Immediate test: *M* = .13, *SD* = .17, *SEM* = .02; 1-week retention test: *M* = .20, *SD* = .21, *SEM* = .02). More importantly, a significant main effect of phonotactic probability was observed, $F(1, 76) = 13.12$, $p < .01$, $\eta_p^2 = .147$, with responses being more accurate to rare (*M* = .18, *SD* = .21, *SEM* = .02) than common sound sequences (*M* = .11, *SD* = ,17, *SEM* = .02). However, these main effects were qualified by a significant phonotactic probability x time x age interaction, $F(2, 76) = 3.417$, $p < .05$, $\eta_p^2 = .083$. The interaction between time and age also was significant, $F(2, 76) = 3.41$, $p < .05$, $\eta_p^2 = .082$. To further examine these interactions, the effect of phonotactic probability and age were examined at each test point. At the immediate test point, significant main effects of age, $F(2, 76) = 4.52$, $p < .05$, $\eta_p^2 = .106$ and phonotactic probability, $F(1, 76) = 10.49$, $p < .01$, $\eta_p^2 = .121$ were qualified

by a significant interaction between phonotactic probability and age, $F(2, 76) = 4.26$, $p < .05$, $\eta_p^2 = .101$. As shown in Table 3, 5-year-olds showed significantly higher proportion correct for rare (*M* = .26, *SD* = .23, *SEM* = .05) sound sequences than common (*M* = .10, *SD* = .12, *SEM* = .03), $t = 3.185$, $p < .01$. In contrast, 3-year-olds (Rare: *M* = .12, *SD* = .20, *SEM* = .04; Common: *M* = .07, *SD* = .14; *SEM* = .03) and 4-year-olds (Rare: *M* = .08, *SD* = .12; *SEM* = .02; Common: *M* = .08, *SD* = .15, *SEM* = .03) showed minimal differences in the proportion correct for common versus rare nonwords *t*s < 1.6, $p > .12$. At the 1-week retention test point, the main effect of phonotactic probability was significant, $F(1, 76) = 6.18$, $p < .05$, $\eta_p^2 = .076$. Here, as shown in Table 3, 3-, 4-, and 5-year olds showed a significantly higher proportion correct for rare sound sequences (*M* = .22, *SD* = .21, *SEM* = .02) than common (*M* = .15, *SD* = .19, *SEM* = .02) sound sequences.

The second follow up ANOVA to examine the effect of phonotactic probability for dense nonwords showed only the main effect of time as significant, $F(1, 76) = 18.39$, $p < .001$, $\eta_p^2 = .195$, with the direction of the effect mirroring the main ANOVA (Immediate test point: *M* = .15, *SD* = .18, *SEM* = .02; 1-week retention test point: *M* = .22, *SD* = .22, *SEM* = .02). Thus, the proportion correct for common (*M* = .18, *SD* = .20, *SEM* = .02) and rare sound sequences (*M* = .18, *SD* = .20, *SEM* = .02) was similar across both test points and across the three age groups for dense nonwords.

The third follow up ANOVA to examine the effect of neighborhood density for rare sound sequences showed a significant main effect of time, $F(1, 76) = 17.33$, $p < .001$, $\eta_p^2 = .186$, mirroring that observed in the main ANOVA (Immediate test point: *M* = .15, *SD* = .18; *SEM* = .01; 1-week retention test point: *M* = .22, *SD* = .21, *SEM* = .02). This was qualified by a significant interaction between neighborhood density, time, and age, $F(2, 76) = 5.23$, $p < .01$, η_p^2

= .121. To examine this interaction, the effects of neighborhood density and age were examined at each time point. At the immediate test point, only the main effect of age was significant, $F(2, 76) = 4.69$, $p < .05$, $\eta_p^2 = .110$, with a significantly greater proportion correct for all nonwords by 5-year-olds ($M = .21$, $SD = .20$, $SEM = .04$) compared to 3- ($M = .12$, $SD = .19$, $SEM = .04$), $t(44) = 2.2$, $p < .01$, and 4-year olds ($M = .12$, $SD = .15$, $SEM = .03$), $t(54) = 3.1$, $p < .01$. Turning to the 1-week retention test point, no main effects or interactions were significant. Taken together, the proportion correct for dense ($M = .18$, $SD = .20$, $SEM = .02$) and sparse nonwords ($M = .18$, $SD = .21$, $SEM = .02$) was similar across both test points and across age groups (see Table 3).

The fourth follow up ANOVA to examine the effect of neighborhood density for common sound sequences showed a significant main effect of time, $F(1, 76) = 15.13$, $p < .001$, $\eta_p^2 = .166$, similar to the main ANOVA (Immediate test point $M = .11$, $SD = .19$, $SEM = .02$; 1-week retention test point: $M = .18$, $SD = .21$, $SEM = .02$). The main effect of neighborhood density also was significant, $F(1, 76) = 12.41$, $p < .01$, $\eta_p^2 = .140$ with a greater proportion correct for dense nonwords ($M = .18$, $SD = .20$, $SEM = .02$) than sparse ($M = .11$, $SD = .17$, $SEM = .01$) nonwords at both test points and at all ages.

To ensure that the pattern of results for each phonotactic probability/neighborhood density condition converged with the results from the statistical analyses across the four semantic categories, item data were inspected. The proportion correct for individual nonwords (collapsed across participants) was visually inspected to determine whether the majority of items followed the pattern reported for the participant analysis (collapsed across items). Additionally, difference scores for each semantic category were calculated to show whether or not the items in each semantic category converged with the participant analysis (i.e., rare sound sequence advantage

for sparse nonwords and dense advantage for common sound sequences). In most cases, this descriptive analysis of item data converged with the participant analysis. In other words, across the majority of semantic categories, the proportion correct was consistently the highest for the nonword with rare sound sequences and a sparse neighborhood and for the nonword with common sound sequences and a dense neighborhood. The pattern of effects for item data within semantic category is shown in Appendix A.

Experiment 1 Summary

Experiment 1 contrasted nonwords with common and rare sound sequences with one another *within* the same story while also contrasting dense and sparse nonwords against one another *across* different stories. Results showed that rare sound sequences were learned significantly better than common sound sequences, but only for sparse nonwords. Likewise, dense nonwords were learned significantly better than sparse nonwords, but only for common sound sequences.

The developmental effects of phonotactic probability observed in this experiment interacted with time (i.e., immediate/early learning time point vs. retention/late learning time point). Specifically, the rare sound sequence advantage for sparse nonwords was robust at the early and later time points of learning (i.e., immediate and retention) for 5-year-olds, but was only observed at the later time point (i.e., 1-week retention) for 3- and 4-year-olds. Similar interactions with time and age were not observed for neighborhood density effects. Specifically, the dense advantage for common sound sequences was similar at both time points and it was robust across the three ages examined.

Results from Experiment 1 alone are not adequate to address whether or not the effects of phonotactic probability and neighborhood density vary depending on context. The difference

between the story context in Experiment 1 and Experiment 2 was designed to answer this question and so it will be addressed following Experiment 2.

Experiment 2

Experiment 2 contrasted nonwords against one another *within* a story so that one nonword was dense and the other was sparse (i.e., neighborhood density). This pairing also resulted in contrasting common and rare nonwords *across* stories (i.e., phonotactic probability).

Methodology

Participants. The recruitment and testing procedures in Experiment 2 were the same as in Experiment 1. The mean, standard deviation and range of standardized test scores for each age group are shown in Table 1. Three groups of typically developing children were recruited for this study: twenty-three 3-year-olds (8 boys and 15 girls*)*, thirty-two 4-year-olds (20 boys and 12 girls*)*, and thirty-one 5-year-olds (12 boys and 19 girls). All children correctly produced the sounds used in the nonword stimuli both in real words and in the imitation of the nonword stimuli. None of the children in Experiment 2 participated in Experiment 1. A comparison of participants across Experiment 1 and Experiment 2 showed no significant difference in standardized test scores, all Fs < .846, all ps > .619, or in overall performance on the experimental word learning tasks, all Fs < .1.23, all ps > .379.

Materials and procedures

The same nonword stimuli, novel object referents and story scripts, measures of learning and procedures used in Experiment 1 were used for Experiment 2. The only difference between Experiment 1 and Experiment 2 is that within each semantic category set within the same story (i.e., toys, pets, horns, and candy machines in story 1 and story 2) the two nonwords were paired so that phonotactic probability was held constant while varying neighborhood density (e.g., Toy

1 paired with a common/dense nonword vs. Toy 2 paired with a common/sparse nonword). The pairing for Experiment 2 is directly opposite the pairing for Experiment 1. Thus, in complement to Experiment 1, each semantic category (e.g., toys) in the same story contrasted a common/dense nonword with a common/sparse nonword or a rare/sparse nonword with a rare/dense nonword. In turn, this pairing contrasted a semantic category in one story where both nonwords were composed of rare sound sequences against the same semantic category in the second story where both nonwords were composed of common sound sequences (e.g., Toys 1 and 2 in story 1 had common sound sequences but Toys 3 and 4 in story 2 had rare sound sequences). This assignment allowed for two comparisons: 1) dense versus sparse (neighborhood density) within a semantic category and within the same story (e.g., Toy 1 in story 1 was paired with a rare/dense nonword and Toy 2 in story 1 was paired with a rare/sparse nonword) and 2) common versus rare sound sequences within a semantic category, but across a story (e.g., both toys in story 1 were paired with nonwords that have rare sound sequences and both toys in story 2 were paired with nonwords that have common sound sequences). This procedure for pairing nonwords with novel object referents within and across stories is the critical difference between Experiment 1 and Experiment 2. Each of the eight nonword-novel object referent pairs was embedded into one of two stories following these procedures with the order of presentation of the stories counterbalanced across participants.

Reliability

Primary judges and reliability judges were the same across Experiments 1 and 2 for transcription, scoring, and procedural reliability. Transcription reliability for consonants in nonword stimuli and real words on the GFTA-2 was computed for 20% of the sample in Experiment 2. Inter-judge transcription reliability for real words was 97% (*SD =2%*, range =

94% to 100%). Inter-judge transcription reliability for nonwords was 97% (*SD = 3%*, range = 88% to 100%).

Scoring reliability was computed for 21% of the sample following the same procedures used for Experiment 1. Inter-judge scoring reliability for the nonword score was 98% (*SD = 2%*, range = 93% - 100%). Inter-judge scoring reliability for the correct score was 98% (*SD = 3%*, range = 93% -100%).

Data collection procedural reliability was computed for 24% of the sample. Inter-judge procedural reliability was 94% (*SD* = 5%, range = 84% - 100%).

Statistical Analyses

Similar to Experiment 1, the dependent variable in Experiment 2 was the proportion of nonwords correct in the picture naming task (i.e., two to three of three phonemes) for each phonotactic probability/neighborhood density condition (i.e., common/dense, common/sparse, rare/dense, and rare/sparse) at two time points (i.e., immediately after exposure and 1-week after exposure) for each age (i.e., 3-, 4-, and 5-years). The exact same analysis procedures used in Experiment 1 were used for Experiment 2. Therefore, all significant multi-way interactions were examined in the same manner as Experiment 1. To be consistent with Experiment 1, significant main effects of time and significant interactions between time and age will be reported, but will not be explicitly discussed because they are not directly relevant to the research questions. As in Experiment 1, non-significant effects in the follow-up analyses will not be explicitly reported, all Fs < 3.78, all ps > .05, all η_p^2s < .07.

Results

Table 4 presents the raw accuracy data (i.e., means, standard deviations, and 95% confidence intervals) for all four experimental conditions (i.e., common-dense, common-sparse,

rare-dense, rare-sparse) at each time point (i.e., immediate vs. retention) for each age (i.e., 3-, 4-, and 5- years).

The main 2 (phonotactic probability) x 2 (neighborhood density) x 2 (time) x 3 (age) ANOVA showed significant main effects of time $F(1, 83) = 42.80$, $p < .001$, $\eta_p^2 = .340$, with higher accuracy at the 1-week retention test point ($M = .18$, $SD = .22$, $SEM = .02$) compared to the immediate test point ($M = .10$, $SD = .15$, $SEM = .01$). The main effect of phonotactic probability was also significant, $F(1, 83) = 9.75$, $p < .01$, $\eta_p^2 = .105$, with higher accuracy for rare ($M = .16$, $SD = .21$, $SEM = .02$) compared to common sound sequences ($M = .12$, $SD = .17$, $SEM = .02$). The main effect of neighborhood density was not significant, $F(1, 83) = 1.25$, $p = .267$, $\eta_p^2 = .015$, with similar accuracy for dense ($M = .13$, $SD = .19$, $SEM = .01$) and sparse ($M = .15$, $SD = .20$, $SEM = .01$) nonwords. Likewise, the main effect of age was not significant, $F(2, 83) = 1.44$, $p = .243$, $\eta_p^2 = .034$, with similar accuracy observed across 3-year-olds ($M = .16$, $SD = .21$, $SEM = .02$), 4-year-olds ($M = .11$, $SD = .17$, $SEM = .02$) and 5-year-olds ($M = .16$, $SD = .19$, $SEM = .02$). Significant two-way interactions between time and age, $F(2, 83) = 3.54$, $p < .05$, $\eta_p^2 = .079$ and between neighborhood density and phonotactic probability, $F(1, 83) = 4.69$, $p < .05$, $\eta_p^2 = .053$ were qualified by a significant 4-way interaction between phonotactic probability, neighborhood density, time, and age $F(2, 83) = 6.88$, $p < .01$, $\eta_p^2 = .142$. None of the remaining two- and three-way interactions involving neighborhood density, phonotactic probability, time and age were significant, all Fs < 1.94, all ps > .168, all η_p^2s < .036. Follow-up ANOVAs were conducted as previously described to determine the pattern of effects of phonotactic probability for dense and sparse nonwords across ages and time and to determine the pattern of effects of neighborhood density for common and rare sound sequences across ages and time.

The first follow up ANOVA to examine the effect of phonotactic probability for sparse nonwords at each test point and age showed significant main effects of time $F(1, 83) = 23.74$, $p < .001$, $\eta_p^2 = .222$ and phonotactic probability $F(1, 83) = 12.03$, $p < .01$, $\eta_p^2 = .127$. The main effect of time was consistent with the main ANOVA (Immediate test point *M* = .11, *SD* = .16, *SEM* = .02; 1-week retention test point: *M* = .19, *SD* = .22, *SEM* = .02;). In terms of the main effect of phonotactic probability, the proportion correct for rare sound sequences (*M* = .19, *SD* = .22, *SEM* = .02) was significantly greater than proportion correct for common sound sequences (*M* = .11, *SD* = .16, *SEM* = .02) for all ages at both test points. Although not relevant to the research questions, there was a significant interaction between time and age, $F(2, 83) = 4.69$, $p < .05$, $\eta_p^2 = .101$, which was not analyzed further.

The second follow up ANOVA to examine the effect of phonotactic probability for dense nonwords at each test point and age showed a significant main effect of time, $F(1, 83) = 30.77$, $p < .001$, $\eta_p^2 = .270$ that mirrored the main ANOVA (Immediate test point: *M* = .09, *SD* = .13, *SEM* = .01; 1-week retention test point: *M* = .18, *SD* = .22, *SEM* = .02). This was qualified by a significant interaction between phonotactic probability, age, and time $F(2, 83) = 6.6$, $p < .01$, $\eta_p^2 = .137$. To explore this interaction, the effect of phonotactic probability and age was examined at each test point. At the immediate test point point, the interaction between phonotactic probability and age was significant $F(2, 83) = 3.54$, $p < .05$, $\eta_p^2 = .08$, however follow up comparisons of phonotactic probability were not statistically significant for any of the age groups, all *p*s > .184. At the 1-week retention test point, neither the main effect of phonotactic probability nor the interaction between phonotactic probability and age were significant *Fs* < 1.6, *ps* >.20. Thus, as shown in Table 4, similar performance was observed for common (*M* = .13, *SD* = .18, *SEM* =

.02) and rare (*M* = .13, *SD* = .19, *SEM* = .01) sound sequences for dense nonwords for all ages at both time points.

The third follow up ANOVA to examine the effect of neighborhood density for rare sound sequences at each test point and age showed significant main effects of time, $F(1, 83) = 28.37$, $p < ,001$, $\eta_p^2 = .255$ and neighborhood density, $F(1, 83) = 4.64$, $p < .05$, $\eta_p^2 = .053$. The main effect of time mirrored the main ANOVA (Immediate test point: *M* = .12, *SD* = .16, *SEM* = .01; 1-week retention test point: *M* = .20, *SD* = .24, *SEM* = .02). Turning to the main effect of neighborhood density, proportion correct for sparse nonwords (*M* = .19, *SD* = .22, *SEM* = .02) was significantly higher than proportion correct for dense nonwords (*M* = .13, *SD* = .19, *SEM* = .02). However, these main effects were qualified by a significant interaction between neighborhood density, time and age $F(2, 83) = 6.17$, $p < .01$, $\eta_p^2 = .129$. Therefore, the effect of neighborhood density and age was examined at each time point. At the immediate test point, only the interaction between neighborhood density and age was significant, $F(2, 83) = 5.64$, $p < .01$, $\eta_p^2 = .120$. Table 4 shows that for 5-year-olds, proportion correct for sparse nonwords (*M* = .22, *SD* = .20, *SEM* = .04) was significantly greater than proportion correct for dense nonwords (*M* = .07, *SD* = .12, *SEM* = .02), $t = 3.649$, $p < .01$. In contrast, no significant effect of neighborhood density was observed for 3-year-olds (Sparse: *M* = .08, *SD* = .14, *SEM* = .03; Dense: *M* = .13, *SD* = .15, *SEM* =.03*)* or for 4-year-olds (Sparse: *M* = .13, *SD* = .17, *SEM* = .04; Dense: *M* = .09, *SD* = .14, *SEM* = .03), *t*s < 1.359, *p*s > .08. At the 1-week retention test point, no main effects or interactions were significant. Here, proportion correct for dense nonwords (*M* = .17, *SD* = .22, *SEM* = .02) was similar to proportion correct for sparse nonwords (*M* = .23, *SD* = .25, *SEM* = .03) for each age group (see Table 4).

The fourth follow up ANOVA to examine the effect of neighborhood density for common sound sequences at each test point and for all ages showed that only the effect of time was significant, $F(1, 83) = 31$, $p < .001$, $\eta_p^2 = .272$, such that all children were more accurate at the 1-week retention test point ($M = .16$, $SD = .20$, $SEM = .02$) than at the immediate test point ($M = .08$, $SD = .14$, SEM .01). No other main effects or interactions were significant. Thus, for common sound sequences, proportion correct for dense nonwords ($M = .13$, $SD = .18$, $SEM = .02$) was similar to proportion correct for sparse nonwords ($M = .11$, $SD = .16$, $SEM = .02$).

To ensure that the pattern of results for each phonotactic probability/neighborhood density condition converged with the results from the statistical analyses across the four semantic categories, item data from Experiment 2 were inspected using the same technique reported for Experiment 1 (i.e., difference score calculations). Like Experiment 1, in the majority of cases, this descriptive analysis of item data converged with the participant analysis. Across the majority of semantic categories, the proportion correct was consistently the highest for the nonword with rare sound sequences and a sparse neighborhood. The pattern of effects for item data within semantic category is shown in Appendix B.

Experiment 2 Summary

Experiment 2 contrasted common versus rare sound sequences against one another *across* different stories and dense versus sparse nonwords against one another *within* the same story. All children learned rare sound sequences significantly better than common sound sequences, but only for sparse nonwords. This rare/sparse advantage was consistent across the immediate and 1-week retention time points. In terms of the effect of neighborhood density, sparse nonwords were learned significantly better than dense nonwords, but only for rare sound sequences by 5-year-olds and only at the immediate time point. This effect of neighborhood density for 5-year-olds

was the only developmental difference observed in this experiment and the only evidence of a difference observed for early versus late learning components.

The issue of whether or not the effects of phonotactic probability and neighborhood density vary depending on context was addressed by comparing the results across Experiment 1 and Experiment 2. This was examined by a 2 (phonotactic probability) x 2 (neighborhood density) x 2 (time) x 3 (age) x 2 (context) ANOVA. A significant main effect of phonotactic probability was obtained, $F(1, 159) = 15.8$, $p < .001$, $\eta_p^2 = .09$, with higher accuracy for rare (*M* = .17, *SD* = .20, *SEM* = .01) than common sound sequences (*M* = .14, *SD* = .18, *SEM* = .01). Importantly, there were no significant interactions involving phonotactic probability and context, all Fs < 2.0, all ps > .155, all $\eta_p^2 < .02$. Thus, the effect of phonotactic probability on word learning was similar across Experiment 1 and 2 with higher accuracy for rare than common sound sequences in both contexts.

In contrast, the previously described ANOVA showed a significant interaction between neighborhood density and context, $F(1, 159) = 5.38$, $p < .05$, $\eta_p^2 = .033$. The results reported for Experiment 1 and 2 provide the follow-up data to understand this interaction. Specifically, as previously described, there was a significant main effect of neighborhood density in Experiment 1 with higher accuracy for dense (*M = .18, SD = .20, SEM = .02*) than sparse neighborhoods (*M = .15, SD* = .19, *SEM* = .02), although this was qualified by an interaction with phonotactic probability. In contrast, there was no significant main effect of neighborhood density in Experiment 2 with similar accuracy across dense (*M* = .13, *SD* = .19, *SEM* = .02) and sparse neighborhoods (*M* = .15, *SD* = .20, *SEM* = .02). When effects of neighborhood density did arise in Experiment 2 (i.e., for 5-year-olds at immediate test), the direction of the effect was opposite

of that found in Experiment 1, with higher accuracy for sparse than dense neighborhoods. Taken together, the effect of neighborhood density varied by context.

General Discussion and Summary

The current studies were designed to address gaps in the literature on the role of form characteristics in word learning by preschool children. Also of interest was whether or not phonotactic probability and neighborhood density affect different components of learning (i.e., triggering vs. lexical configuration and engagement) and whether this differed by exposure context. The pattern of interactions observed in this study highlights the complex nature of word learning across development and raises important differences between children and adult's word learning. The current studies yield interactions between phonotactic probability and neighborhood density in word learning by preschool children. This contrasts with previous findings from studies of adult word learning (Storkel et al., 2006), which failed to find an interaction between phonotactic probability and neighborhood density. It appears that preschool children, unlike adults, benefit from a convergence of form characteristics when learning new words.

Across the current studies, children learned rare sound sequences from sparse neighborhoods significantly better than common sound sequences from sparse neighborhoods. The effect of phonotactic probability was consistent with that of the previous adult word learning study. In that study, Storkel et al. (2006) hypothesized the rare sound sequence advantage to reflect a listener's ability to more easily identify unique sounding words as novel thereby more rapidly triggering the process of learning. Since learning is initiated sooner for rare sound sequences, it is likely that fewer exposures to nonwords with rare sound sequences are needed as opposed to nonwords with common sound sequences. More confusion is likely to occur when

learning nonwords composed of common sound sequences because these nonwords are similar to other known words in the lexicon (Frisch et al., 2000; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). Because more exposure to words with common sound sequences may be required to reconcile the similarity with other known words, the onset of learning may be delayed in comparison to words composed of rare sound sequences (Storkel et al., 2006).

Crucially, the effect of phonotactic probability for adult word learning was not dependent on the neighborhood density of the novel word, whereas the effect for preschool word learning was dependent on the neighborhood density of the novel words. This suggests that neighborhood density, and by extension lexical representations, may play a role in triggering word learning by preschool children, but not by adults. That is, when a learner is initially presented with a novel word, existing lexical representations will also be activated (Gaskell & Dumay, 2003). If no existing lexical representation sufficiently matches the novel word, then this would provide an additional indication that the word is novel and a new lexical representation needs to be created (i.e., triggering learning). One difference between the learning of sparse and dense novel words is in the number of existing lexical representations that are activated during exposure. When a learner encounters a novel sparse word fewer existing lexical representations are activated than for a dense word. This initial activation of few existing lexical representations may speed detection of the mismatch between the input and existing lexical representations, thereby more efficiently triggering word learning than in the case where many existing lexical representations are activated, as in a dense neighborhood. In this way, sparse neighborhoods and rare sound sequences converge to signal that a word is novel. This condition is optimal for the learner because the word's distinctiveness triggers word learning more rapidly than in other conditions (e.g., common/sparse).

Turning to neighborhood density, Experiment 1 demonstrated that preschool children learn common sound sequences from dense neighborhoods more readily than common sound sequences from sparse neighborhoods. This effect of density is consistent with the past study of adult word learning, which hypothesized that neighborhood density influenced lexical configuration and engagement (Storkel et al., 2006). Specifically, dense words are held in working memory better than sparse words (Roodenrys & Hinton, 2002; Thomson et al., 2005; Thorn & Frankish, 2005), potentially supporting the creation of an accurate and/or detailed new lexical representation for dense words when compared to sparse words. In addition, dense neighborhoods could entail an advantage over sparse neighborhoods during configuration and engagement. That is, a new lexical representation in a dense neighborhood would form many links with existing lexical representations. As a result, these multiple links may serve to reinforce or strengthen the new lexical representation.

As with phonotactic probability, the effect of neighborhood density on word learning by preschool children was dependent on the phonotactic probability of the novel sound sequence, indicating the benefit of converging form characteristics for configuration and engagement. Specifically, common sound sequences are held in working memory better than rare sound sequences (Gathercole et al., 1999). Thus, common sound sequences and dense neighborhoods converge to create optimal working memory support, facilitating the creation of an accurate and detailed new lexical representation when compared to other conditions. Likewise, common sound sequences also could play a role in engagement. To this point, discussion of engagement has focused on the integration of a new lexical representation with existing lexical representations. However, engagement also entails the integration of the new lexical representation with representations of other types, in this case phonological representations. That

is, the new lexical representation must form links with its component phonological representations. Common sound sequences are hypothesized to have more stable and robust representations (Munson, Swenson, & Manthei, 2005) and greater stored activation (Vitevitch & Luce, 1998, 1999) than rare sound sequences. As a result, creation of links between a new lexical representation and common phonological representations may strengthen the new lexical representation. Taken together, dense neighborhoods and common sound sequences converge to create optimal conditions for configuration and/or engagement during word learning by preschool children.

Interestingly, this convergence of dense neighborhoods and common sound sequences was only observed when the density manipulation was implemented across stories (i.e., Experiment 1). In contrast, when the density manipulation was implemented within stories (i.e., Experiment 2) a different convergence of form characteristics arose as optimal, although this was only observed for older children at one time point. Specifically, older children learned rare sound sequences in sparse neighborhoods more readily than rare sound sequences in dense neighborhoods. This pattern is consistent with a competitive learning environment and is similar to the results of Hollich et al. (2002) and Swingley & Aslin (2007). This could arise through word recognition processes during repeated exposure to novel words. In most models of word recognition, similar lexical representations either have inhibitory connections to one another (e.g., Auer, 1993; McClelland & Elman, 1986) or they have no direct connection to one another, competing for selection through the activation and selection process (e.g., Luce & Pisoni, 1998; Norris, 1994). Thus, during repeated exposure to novel words, existing lexical representations would either inhibit the newly created lexical representation or would compete for selection with the newly created lexical representation, thereby impeding configuration. The extent of this

inhibition or competition would depend on the number of similar existing lexical representations. Specifically, greater inhibition or competition would be encountered in dense neighborhoods than in sparse neighborhoods. Likewise, phonotactic probability could increase the amount of inhibition or competition due to the interaction between phonological and lexical representations. Specifically, activation spreads back and forth between phonological and lexical representations. The amount of activation spread from the phonological representations back to the lexical representations will be influenced by the stored activation in the phonological representation, which is affected by phonotactic probability. That is, common sound sequences will send greater activation back to lexical representations than rare sound sequences will, leading to greater inhibition or competition. Taken together, decreased inhibition or competition during word recognition, as would occur for sparse neighborhoods and rare sound sequences, would enhance configuration.

Why would pairing words within a story lead to the competitive learning environment just described, whereas pairing words across stories would not? For the within story manipulation of neighborhood density, the asymmetry in the number of lexical competitors is immediately present during learning. That is, one novel word is sparse, with few competitors, and one novel word is dense, with many competitors. In contrast, for the across story manipulation of neighborhood density, the asymmetry in the number of lexical competitors is not immediately present during learning. On a given learning trial, the two novel words the child is exposed to have a similar number of lexical competitors (i.e., matched neighborhood density). In this way, the within story manipulation of neighborhood density emphasizes the inhibition or competition between lexical representations during learning, whereas the across story manipulation of neighborhood density does not. This suggests that learning certain types of

words can be more challenging depending on the context in which they are presented to the learner. Specifically, dense words are learned more easily in a non-competitive context whereas sparse words are learned more easily in a competitive context. This may inform theories addressing how children add new words to their existing vocabulary.

In terms of developmental changes in word learning, no consistent developmental differences were observed within the narrow span of preschool ages tested here, and those that were detected were further complicated by interactions with time. Specifically, only 5-year-olds in Experiment 1 learned novel words with rare sound sequences in sparse neighborhoods more readily than novel words with common sound sequences in sparse neighborhoods at the immediate test time, although children of all ages showed this same pattern at the 1-week retention time. Likewise, only 5-year-olds in Experiment 2 learned novel words with rare sound sequences in sparse neighborhoods more readily than novel words with rare sound sequences in dense neighborhoods at the immediate test time, and this effect was no longer significant for any age group at the retention test. Taken together, developmental differences across the preschool period appear to relate more to the timing of effects rather than to the direction of effects. One possible developmental effect that might be considered is the acquisition of the ability to detect the natural distribution of correlations between phonotactic probability and neighborhood density that are apparent in the language. Recall that phonotactic probability and neighborhood density are correlated so that sparse words tend to be composed of rare sound sequences and dense words tend to be composed of common sound sequences (Storkel, 2004b; Vitevitch et al., 1999). The fact that only the 5-year olds showed evidence of learning nonwords in the correlated condition at the immediate test time in Experiment 1 (i.e., common/dense) and at all in Experiment 2 (i.e., rare/sparse at immediate time point only) might reflect a developmental effect

in the ability to recognize the correlation between these two form characteristics in the language. This explanation needs to be further delineated in future studies to determine whether or not there is a developmental component to detecting these natural correlations or whether this finding was purely related to timing effects in this study.

In contrast, the most striking developmental differences in the role of phonotactic probability and neighborhood density are those that arise between the preschool children of this study and the adults of the previous study (Storkel et al., 2006). Specifically, preschool children appear to benefit from a convergence of phonotactic probability and neighborhood density during word learning. In contrast, the past study of adult word learning yielded no significant interaction between phonotactic probability and neighborhood density, suggesting that adults may be less reliant on a convergence of form characteristics to support word learning. With development, it is possible that the weighting of phonotactic probability and neighborhood density for each component of word learning changes, as would be predicted by the Emergentist Coalition Model (Hirsh-Pasek et al., 2000). Presumably, with age, phonotactic probability may be weighted more strongly than neighborhood density for triggering word learning and neighborhood density may be weighted more strongly than phonotactic probability for lexical configuration and engagement as observed by main effects, but no interactions (Storkel, in press).

Why would this re-weighting occur? Presumably, different representations may become more closely aligned with specific components of the word learning process to increase the efficiency of word learning. That is, phonological representations would be primarily involved in triggering word learning while lexical representations would be primarily involved in configuration and engagement (Storkel, in press). The interactive convergence of rare sound

sequences and sparse neighborhoods observed in children may also reflect the developmental nature of representations for children. Less robust phonological and lexical representations in children may require a convergence of form characteristics that is not necessary for adults because their lexical and phonological representations are fully specified. Before representations are robust and detailed, children's phonological and lexical representations may be dependent on one another which would explain the convergence of form characteristics observed in this study.

Conclusion

The combined results of Experiments 1 and 2 provide new evidence on the role of form characteristics in word learning by preschoolers and the extent to which this use is related to age, word learning component (i.e., immediate learning versus later learning/retention), and exposure context. The findings from the current experiments suggest that unlike adults, preschool children benefit from a convergence of phonotactic probability and neighborhood density for word learning. Specifically, rare sound sequences and sparse neighborhoods converge to efficiently trigger word learning. In contrast, optimal convergence for lexical configuration and engagement is dependent on exposure context. In particular, common sound sequences and dense neighborhoods converge to facilitate configuration and/or engagement when density is manipulated across stories, whereas rare sound sequences and sparse neighborhoods converge to facilitate configuration when density is manipulated within the same story. The differences between this study and a previous study of adult word learning (Storkel et al., 2006) suggest that a re-weighting of form characteristics may occur sometime after preschool, but before adulthood. Specifically, children's phonological and lexical representations may be more dependent on one another resulting in a convergence of characteristics for optimal word learning that is not observed by adults.

References

ASHA. (1997). Guidelines for audiologic screening. *Asha*, IV-74a - IV-74i.

Aslin, R. N., & Swingley, D. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*, 99-132.

Auer, E. T. (1993). *Dynamic processing in spoken word recognition: The influence of paradigmatic and syntagmatic states.* Unpublished Dissertation, State University of New York at Buffalo, Buffalo, NY.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language, 44*, 568-591.

Brownell, R. (2000a). *Expressive one-word picture vocabulary test - 3rd edition*. Novato, CA: Academic Therapy Publications.

Brownell, R. (2000b). *Receptive one-word picture vocabulary test - 2nd edition*. Novato, CA: Academic Therapy Publications.

DeBrunhoff, L. (1981). *Babar's anniversary album*. New York, NY: Random House.

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language, 42*, 481-496.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105-132.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 84-95.

Geisel, T. S., & Geisel, A. S. (1954). *Horton hears a who!* New York: Random House.

Geisel, T. S., & Geisel, A. S. (1958). *Cat in the hat comes back*. New York: Random House.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2*. Circles Pines, MN: American Guidance Service.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language. Special Issue: Computer models of impaired language, 59,* 267-333.

Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An emergentist coalition model for word learning: Mapping words to objects is a product of the interaction of multiple cues. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. Woodward, N. Akhtar, M. Tomasello & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 136-164). Oxford: Oxford University Press.

Hollich, G., Jusczyk, P. W., & Luce, P. A. (2002). Lexical neighborhood effects in 17-month-old word learning. In B. Skarabela, S. Fish & A. H.-J. Do (Eds.), *Proceedings of the 26th annual Boston University Conference on Language Development* (Vol. 1, pp. 314-323). Sommerville, MA: Cascadilla.

Jarvis, B. G. (2002). DirectRT research software (Version 2002). New York, NY: Empirisoft.

Kolson, C. J. (1960). *The vocabulary of kindergarten children.* Unpublished Doctoral Dissertation, University of Pittsburgh, Pittsburgh.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55,* 306-353.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing, 19*, 1-36.

Mayer, M. (1992). *Professor wormbog in search for the zipperump-a-zoo*. Italy: Rainbird Press.

Mayer, M. (1993). *Little critter's read-it-yourself storybook: Six funny easy-to-read stories*. New York: Golden Book.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

Moe, A. J., Hopkins, K. J., & Rush, R. T. (1982). *The vocabulary of first grade children*. Springfield, IL: Thomas.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research, 47*, 1048-1058.

Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: Evidence from repetition tasks. *Journal of Speech, Language and Hearing Research, 48*, 108-124.

Newman, R., Samuelson, L., & Gupta, P. (2008). *Learning novel neighbors: Distributed mappings help children and connectionist models.* Paper presented at the 30th Annual Meeting of the Cognitive Science Society, Washington D.C.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon. In *Research on Spoken Language Processing Report No. 10* (pp. 357-376). Bloomington, IN: Speech Research Laboratory, Indiana University.

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 29-33.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321-1337.

Storkel, H. L. (2003). Learning new words II: Phonotactic probability in verb learning. *Journal of Speech, Language, and Hearing Research, 46,* 1312-1323.

Storkel, H. L. (2004a). The emerging lexicon of children with phonological delays: Phonotactic constraints and probability in acquisition. *Journal of Speech, Language, and Hearing Research, 47,* 1194-1212.

Storkel, H. L. (2004b). Methods for Minimizing the Confounding Effects of Word Length in the Analysis of Phonotactic Probability and Neighborhood Density. *Journal of Speech, Language, and Hearing Research, 47,* 1454-1468.

Storkel, H. L. (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *J Child Lang, 36,* 291-321.

Storkel, H. L. (in press). The developing mental lexicon of children with specific language impairment. In J. Guendouzi, F. Loncke & M. J. Williams (Eds.), *The handbook of psycholinguistic & cognitive processes: Perspectives in communication disorders* Taylor & Francis.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49,* 1175-1192.

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language, 32*, 827-853.

Storkel, H. L., & Rogers, M. A. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics & Phonetics, 14*, 407-425.

Thomson, J. M., Richardson, U., & Goswami, U. (2005). Phonological similarity neighborhoods and children's short-term memory: typical development and dyslexia. *Memory & Cognition, 33*, 1210-1219.

Thorn, A. S., & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 729-735.

Treiman, R., Fowler, C. A., Gross, J., Berch, D., & Weatherson, S. (1995). Syllable structure or word structure? Evidence for onset and rime units in disyllabic and trisyllabic stimuli. *Journal of Memory and Language, 34*, 132-155.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*, 325-329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory of Language, 40*, 374-408.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech, 40*, 47-62.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language. Special Issue: Mental lexicon, 68,* 306-311.

Ziegler, J. C., & Goswami, U. C. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3-29.

Author Note

Jill R. Hoover, Department of Speech & Hearing Sciences, Indiana University. Holly L. Storkel, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas. Tiffany P. Hogan, Department of Special Education and Communication Disorders, University of Nebraska – Lincoln.

This research was supported by NIH Grants awarded to Holly Storkel (DC04781, DC08095), Jill Hoover (DC009135), and Tiffany Hogan (DC009667). The following NIH grants also supported this research: DC05803, DC000052, and HD02528.

Portions of the research reported here were presented at the 2005 and 2007 conventions for the American Speech-Language-Hearing Association in San Deigo, CA and Boston, MA.

The authors wish to thank the following members of the Word and Sound Learning Laboratory who contributed to stimulus preparation, data entry and data collection: Andy Aschenbrenner, Teresa Brown, Jennie Fox, Andrea Giles, Stephanie Gonzales, Kyung Ae Keum, Jennica Kilwein, Suyeon Lee, Junko Maekawa, Shannon Rogers, Mariza Rosalez, Josie Row, Maki Sueto, Allison Wade, and Courtney Winn.

Correspondence concerning this article should be addressed to Holly Storkel, PhD., Associate Professor, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas, 3001 Dole Human Development Center, 1000 Sunnyside Ave., Lawrence, KS 66045-7555. Email: hstorkel@ku.edu.

Table 1

Means, Standard Deviations, Age Range and Standardized Test Standard Scores for each Group (3- vs. 4- vs. 5- year-olds) in Experiments 1 and 2

	3-year-olds		4-year-olds		5-year-olds	
	Exp. 1	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2
Age						
M	41	43	54	53	63	64
SD	4	3	3	3	2	3
Range	36-47	37-47	48-59	48-59	60-69	60-70
ROWPVT						
M	105	107	104	104	105	109
SD	10	12	13	11	8	11
Range	87-122	81-128	76-131	77-127	90-127	85-130
EOWPVT						
M	106	103	105	100	105	105
SD	7	14	14	12	15	13
Range	93-118	83-135	83-136	68-127	74-129	83-133
GFTA-2						
M	109	109	110	108	110	107
SD	10	7	8	8	4	8
Range	89-123	95-120	92-119	92-120	99-116	89-117

Note. ROWPVT = Receptive One Word Picture Vocabulary Test (*M* = 100; *SD* = 15);

EOWPVT = Expressive One Word Picture Vocabulary Test (*M* = 100; *SD* = 15);

GFTA-2 = Goldman Fristoe Test of Articulation Second Edition (*M* = 100; *SD* = 15).

Table 2

Means, standard deviations, and ranges, respectively, of phonotactic probability and neighborhood density of the Nonword Stimuli

Characteristics based on Adult Corpus

	Common		Rare	
	Dense[1]	Sparse[2]	Dense[3]	Sparse[4]
Positional Segment Frequency	.17	.15	.09	.09
	(.02)	(.03)	(.01)	(.01)
	.15-.19	.11-.20	.08-.10	.08-.10
Biphone Frequency	.006	.007	.002	.002
	(.003)	(.006)	(.000)	(.001)
	.004- .010	.003-.016	.001-.002	.000-.002
Neighborhood Density	16	6	13	7
	(.5)	(1)	(1)	(1)
	15-16	4-7	12-14	6-7

Characteristics based on Child Corpus

	Common		Rare	
	Dense[1]	Sparse[2]	Dense[3]	Sparse[4]
Positional Segment Frequency	.18	.16	.12	.11
	(.02)	(.05)	(.02)	(.02)
	.16-.20	.13-.22	.10-.14	.09-.13

Biphone Frequency	.007	.006	.003	003
	(.002)	(.003)	(.001)	(.002)
	.006- .009	.002-.010	.002-.004	.000-.005
Neighborhood Density	13	6	9	6
	(2)	(3)	(3)	(2)
	10-15	3-9	6-12	3-7

Nonwords by condition:[1]/woʊn jæt nɪd paʊn/ [2]/mɛb paɪb hɑn jaʊn/ [3]/wud jeɪm nʌd haʊd/

[4]/hub tɔɪm nib wʌp/.

Table 3

Means, Standard Deviations, and 95% Confidence Intervals for Condition Accuracy in Experiment 1

	Dense				Sparse			
	Common		Rare		Common		Rare	
	Immediate	Post	Immediate	Post	Immediate	Post	Immediate	Post
3 y/o	.12	.17	.12	.16	.07	.10	.12	.15
	(.20)	(.22)	(.18)	(.19)	(.14)	(.15)	(.20)	(.21)
	[.04, .20]	[.08, .26]	[.05, .19]	[.08, .24]	[.01, .13]	[.04, .16]	[.04, .20]	[.06, .24]
4 y/o	.17	.23	.15	.20	.08	.16	.08	.25
	(.19)	(.22)	(.18)	(.21)	(.15)	(.22)	(.12)	(.20)
	[.11, .23]	[.15, .31]	[.09, .21]	[.13, .27]	[.03, .13]	[.09, .24]	[.04, .12]	[.18, .32]
5 y/o	.15	.24	.16	.28	.10	.18	.26	.24
	(.16)	(.22)	(.16)	(.24)	(.12)	(.17)	(.23)	(.23)
	[.08, .22]	[.15, .33]	[.09, .23]	[.18, .38]	[.05, .15]	[.11, .25]	[.17, .35]	[.15, .33]
All Ages	.15	.22	.14	.21	.08	.15	.15	.22
	(.18)	(.22)	(.17)	(.22)	(.14)	(.19)	(.19)	(.21)

	[.11, .19]	[.17, .27]	[.10, .18]	[.16, .26]	[.05, .11]	[.11, .19]	[.11, .19]	[.17, .27]
All Ages	.18		.18		.11		.18	
& Times	(.20)		(.20)		(.17)		(.21)	
	[.14, .23]		[.13, .22]		[.08, .15]		[.14, .23]	

Note: Standard deviations are noted in round parentheses and the 95% confidence interval is noted in square brackets.

Table 4

Means, Standard Deviations, and 95% Confidence Intervals for Condition Accuracy in Experiment 2

	Dense				Sparse			
	Common		Rare		Common		Rare	
	Immediate	Post	Immediate	Post	Immediate	Post	Immediate	Post
3 y/o	.07	.22	.13	.21	.09	.20	.08	.26
	(.11)	(.27)	(.15)	(.28)	(.14)	(.23)	(.14)	(.26)
	[.02, .11]	[.11, .33]	[.07, .19]	[.1, .32]	[.03, .15]	[.11, .29]	[.02, .14]	[.16, .37]
4 y/o	.06	.16	.09	.10	.08	.09	.13	.20
	(.11)	(.19)	(.14)	(.19)	(.12)	(.15)	(.17)	(.25)
	[.02, .10]	[.10, .23]	[.04, .14]	[.03, .17]	[.04, .12]	[.04, .14]	[.07, .19]	[.11, .29]
5 y/o	.12	.16	.07	.22	.08	.14	.22	.24
	(.17)	(.19)	(.12)	(.20)	(.14)	(.18)	(.20)	(.24)
	[.06, .18]	[.09, .23]	[.03, .11]	[.15, .29]	[.03, .13]	[.08, .20]	[.15, .29]	[.16, .32]
All Ages	.08	.18	.09	.17	.08	.14	.15	.23
	(.14)	(.21)	(.13)	(.22)	(.13)	(.19)	(.18)	(.25)

	[.06, .11]	[.13, .22]	[.06, .12]	[.12, .22]	[.05, .11]	[.10, .18]	[.11, .19]	[.18, .28]
All Ages &	.13		.13		.11		.19	
Times	(.18)		(.19)		(.16)		(.22)	
	[.09, .17]		[.09, .17]		[.07, .14]		[.14, .24]	

Note: Standard deviations are noted in round parentheses and the 95% confidence interval is noted in square brackets.

Appendix A: Mean Difference Scores for each Semantic Category within each Phonotactic Probability/Neighborhood Density Condition in Experiment 1

	[1]Phonotactic Probability Effect: Rare > Common All Ages, Collapsed Across both Times		[2]Neighborhood Density Effect: Dense > Sparse All Ages, Collapsed Across both Times	
	Dense = No	Sparse = Yes	Common = Yes	Rare = No
Toys	.08	-.03	-.07	.04
Horns	-.07	.09	.04	-.12
Candy	-.04	.02	.10	.04
Pets	.02	.14	.18	.06
All Items	0	.07	.06	-.01

Note: [1] Phonotactic Probability Effect: Difference scores were calculated for semantic categories by subtracting the accuracy for common nonwords from the accuracy for rare nonwords collapsed across both time points for all ages. Positive difference scores for sparse nonwords indicate that the item data converged with the participant analysis (i.e., Rare > Common for sparse nonwords). [2] Neighborhood Density Effect: Difference scores were calculated for semantic categories by subtracting the accuracy for sparse nonwords from the accuracy for dense nonwords collapsed across both time points for all ages. Positive difference scores for common sound sequences indicate that the item data converged with the participant analysis (i.e., Dense > Sparse, for common sound sequences).

Appendix B: Mean Difference Scores for each Semantic Category within each Phonotactic Probability/Neighborhood Density Condition in Experiment 2

	[1]Phonotactic Probability Effect: Rare > Common All Ages, Collapsed Across both Times		[2]Neighborhood Density Effect: Sparse > Dense 5-year-olds at immediate time point	
	Dense = No	Sparse = Yes	Common = No	Rare = Yes
Toys	.02	.13	-.07	.26
Horns	-.03	.06	.06	.13
Candy	.18	.10	-.07	.13
Pets	-.15	.04	-.09	.07
All Items	-.02	.06	-.04	.15

Note: [1] Phonotactic Probability Effect: Difference scores were calculated for semantic categories by subtracting the accuracy for common nonwords from the accuracy for rare nonwords collapsed across both time points for all ages. Positive difference scores for sparse nonwords indicate that the item data converged with the participant analysis (i.e., Rare > Common for sparse nonwords). [2] Neighborhood Density Effect: Difference scores were calculated for semantic categories by subtracting the accuracy for dense nonwords from the accuracy for sparse nonwords only at the immediate time point and only for 5-year-olds. Positive difference score for rare sound sequences indicates that the item data converged with the participant analysis (i.e., Sparse > Dense, for rare sound sequences at the immediate time point for 5-year-olds).