NIH Public Access
Author Manuscript
J Speech Lang Hear Res. Author manuscript; available in PMC 2011 August 1.

Published in final edited form as:
J Speech Lang Hear Res. 2010 August ; 53(4): 933–949. doi:10.1044/1092-4388(2009/09-0075).

Differentiating the effects of phonotactic probability and neighborhood density on vocabulary comprehension and production: A comparison of preschool children with versus without phonological delays

Holly L. Storkel, **Junko Maekawa**, and **Jill R. Hoover**
University of Kansas

Abstract

Purpose—The purpose of this study was to differentiate the effect of phonotactic probability from that of neighborhood density on a vocabulary probe administered to preschool children with or without a phonological delay.

Method—Twenty preschool children with functional phonological delays and 34 preschool children with typical language development completed a 121 item vocabulary probe in both an expressive and receptive response format. Words on the vocabulary probe orthogonally varied on phonotactic probability and neighborhood density but were matched on age-of-acquisition, word frequency, word length, semantic set size, concreteness, familiarity, and imagability.

Results—Results showed an interaction between phonotactic probability and neighborhood density with variation across groups. Specifically, the optimal conditions for typically developing children were rare phonotactic probability with sparse neighborhoods and common phonotactic probability with dense neighborhoods. In contrast, only rare phonotactic probability with sparse neighborhoods was optimal for children with phonological delays.

Conclusions—Rare sound sequences and sparse neighborhoods may facilitate triggering of word learning for typically developing children and children with phonological delays. In contrast, common sound sequences and dense neighborhoods may facilitate configuration and engagement for typically developing children but not children with phonological delays due to their weaker phonological and/or lexical representations.

Keywords

word learning; vocabulary; neighborhood density; phonotactic probability; phonological delay

Many models of spoken word recognition, production, and learning assume two types of form representations: phonological and lexical (e.g., Dell, 1988; Gupta & MacWhinney, 1997; Levelt, 1989; Luce, Goldinger, Auer, & Vitevitch, 2000; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; McClelland & Elman, 1986; Norris, 1994). *Phonological representations* correspond to individual sounds with variation across models in the specific unit of sound chosen (e.g., phonetic features, phones, phonemes). *Lexical representations*

Contact author: Holly Storkel, Ph.D., Associate Professor, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas, 3001 Dole Human Development Center, 1000 Sunnyside Avenue, Lawrence, KS 66045-7555. hstorkel@ku.edu..
Holly L. Storkel and Junko Maekawa, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas. Jill R. Hoover, Child Language Doctoral Program, University of Kansas.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

correspond to whole-word sound sequences as an integrated unit. A given word in a language has both a phonological and a lexical representation. For example, the word 'cat' consists of three phonological representations (assuming the phoneme is the sound unit chosen), specifically /k/, /æ/, and /t/, and one lexical representation, specifically /kæt/. As the example illustrates, words will tend to have multiple phonological representations, corresponding to the number of individual sound units in the word, but only one lexical representation.

Two correlated variables have been manipulated in tandem or separately to investigate the influence of phonological and lexical representations on word recognition, production, memory, and learning by adults: phonotactic probability and neighborhood density (e.g., Roodenrys & Hinton, 2002; Storkel, Armbruster, & Hogan, 2006; Thorn & Frankish, 2005; Vitevitch, 2002; Vitevitch, Armbruster, & Chu, 2004; Vitevitch & Luce, 1999). *Phonotactic probability* refers to the likelihood of occurrence of a given sound or pair of sounds in a language and is thought to influence activation of phonological representations. *Neighborhood density* refers to the number of words that differ from a given word by one phoneme and is thought to influence activation of lexical representations.

While phonotactic probability and neighborhood density have received much attention in studies of the fully developed lexicon, mounting evidence suggests that these variables also are relevant in the emerging lexicon of typically developing children. Specifically, phonotactic probability influences speed and/or accuracy of speech production in children from age 2 to adolescence with production of common sound sequences being faster and/or more accurate than production of rare sound sequences, although this effect may be modulated by vocabulary size (e.g., Edwards, Beckman, & Munson, 2004; Munson, Swenson, & Manthei, 2005; Newman & German, 2005; Zamuner, Gerken, & Hammond, 2004). Children also recall more nonwords composed of common sound sequences than those composed of rare sound sequences in working memory tasks (Gathercole, Frankish, Pickering, & Peaker, 1999). In terms of neighborhood density, children recognize words in sparse neighborhoods faster than words in dense neighborhoods, although this effect may be modulated by word frequency and age (Garlock, Walley, & Metsala, 2001; Mainela-Arnold, Evans, & Coady, 2008; Metsala, 1997). Likewise, words in sparse neighborhoods are produced faster and/or more accurately than words in dense neighborhoods, and again this effect may be modulated by age (Munson, Swenson et al., 2005; Newman & German, 2005). In contrast, children recall more words and nonwords from dense neighborhoods than sparse neighborhoods in working memory tasks (Thomson, Richardson, & Goswami, 2005).

Turning to word learning, the majority of studies have tended to examine phonotactic probability and neighborhood density in tandem. These two variables are positively correlated in English such that common sound sequences tend to reside in dense neighborhoods and rare sound sequences tend to reside in sparse neighborhoods (Storkel, 2004c; Vitevitch, Luce, Pisoni, & Auer, 1999). Results from word learning studies manipulating correlated phonotactic probability and neighborhood density have shown that preschool children learn common sound sequences from dense neighborhoods more rapidly than rare sound sequences from sparse neighborhoods (Storkel, 2001, 2003; Storkel & Maekawa, 2005).

Two studies have attempted to differentiate the effect of phonotactic probability from that of neighborhood density in word learning by infants or adults. For the infant study, a database of real words known by infants age 1;4 to 2;6 was analyzed using linear regression so that the contribution of phonotactic probability and neighborhood density to age-of-acquisition could be disentangled (Storkel, 2009). Results showed that infants learned rare sound sequences at earlier ages than common sound sequences. In addition, infants learned words

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in dense neighborhoods at earlier ages than words in sparse neighborhoods (Storkel, 2009). For the adult study, phonotactic probability and neighborhood density were fully crossed when creating the nonwords to be learned (Storkel et al., 2006). Adults were then exposed to these nonwords paired with novel objects and learning was tracked via picture naming. Adults demonstrated the same pattern observed in the infant study, learning rare sound sequences more readily than common sound sequences and learning nonwords in dense neighborhoods more readily than nonwords in sparse neighborhoods (Storkel et al., 2006).

From these findings, it was hypothesized that phonological representations may play a critical role in triggering word learning. Specifically, rare sound sequences may be more rapidly identified as novel than common sound sequences, immediately triggering, and thus speeding, learning. In contrast, lexical representations were hypothesized to play more of a role in configuration, specifically the creation of a new representation in the lexicon, or engagement, namely the integration of a new representation with existing representations (Leach & Samuel, 2007). In terms of configuration, dense neighborhoods have been shown to facilitate maintenance of sound sequences in working memory compared to sparse neighborhoods (Roodenrys & Hinton, 2002; Thomson et al., 2005; Thorn & Frankish, 2005). Consequently, creation of a new lexical representation in long-term memory was presumed to be more successful for a novel word from a dense neighborhood than from a sparse neighborhood because of the greater support from working memory. In terms of engagement, forming connections between a new lexical representation and existing lexical representations may serve to stabilize the new representation. That is, integration of a new lexical representation with many existing representations, as in a dense neighborhood, was assumed to reinforce the new representation more than integration of a new lexical representation with few existing representations, as in a sparse neighborhood.

Due to the hypothesized role of phonological representations in word learning, it is possible that children who have weak phonological representations may show differing effects of phonotactic probability and possibly neighborhood density on word learning. One such group is children with functional phonological delays (Munson, Edwards, & Beckman, 2005a). Children with functional phonological delays experience significant deficits in acquiring the sound system of their native language in the absence of any concomitant deficits in motor, sensory, cognitive, or social abilities (Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986). Studies of word learning in this population have typically examined performance on standardized tests of vocabulary with the results showing that children with phonological delays perform more poorly than their typically developing peers, albeit still within the normal range, and that this vocabulary difference is evident even after the phonological delay has resolved (Felsenfeld, Broen, & McGue, 1992; Shriberg & Kwiatkowski, 1994).

One study did examine the influence of correlated phonotactic probability and neighborhood density on word learning by children with phonological delays (Storkel, 2004b). Results showed that children with phonological delays learned rare sound sequences from sparse neighborhoods more readily than common sound sequences from dense neighborhoods. In contrast, typically developing children showed the opposite pattern, learning common sound sequences from dense neighborhoods more readily than rare sound sequences from sparse neighborhoods. Because only correlated phonotactic probability and neighborhood density were examined, it is difficult to know whether the difference between groups is attributable to differences in the effect of phonotactic probability on triggering word learning, the effect of neighborhood density on word learning configuration and/or engagement, or both. Thus, one goal of the current study was to differentiate the effects of phonotactic probability and neighborhood density on word learning by children with phonological delays to provide evidence of the independent and interactive effects of these two variables in this group of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

children. These data can then be used to form hypotheses about the role of phonotactic probability and neighborhood density in triggering, configuration, and engagement by children with phonological delays.

A final issue is the type of paradigm used to examine the effect of phonotactic probability and neighborhood density on word learning. Specifically, the majority of past studies of phonotactic probability and neighborhood density in word learning have used an experimental word learning paradigm where children were exposed to nonwords paired with novel objects and learning was tracked across exposures (Storkel, 2001, 2003, 2004b; Storkel et al., 2006; Storkel & Maekawa, 2005). This experimental word learning paradigm is somewhat time-consuming to administer, requiring several sessions, and thus may not be practical for clinical or large-scale research applications (e.g., longitudinal studies investigating multiple components of language). However, a few recent studies have suggested that more traditional measures, such as vocabulary checklists or language samples, may provide evidence that converges with findings from experimental word learning paradigms regarding the effects of phonotactic probability and neighborhood density on word learning (Coady & Aslin, 2003; Maekawa & Storkel, 2006; Storkel, 2004a). Moreover, there is emerging evidence that these traditional measures also may be sensitive to the differing effect of each of these variables on distinct components of word learning, namely triggering versus configuration versus engagement (Storkel, 2009). This study seeks to extend these findings to a measure that may be appropriate for a wider age range, specifically an expressive and receptive vocabulary probe modeled after numerous standardized vocabulary tests frequently used in clinical settings and large-scale research studies.

How is it possible that a static vocabulary probe can be sensitive to dynamic word learning processes such as triggering, configuration, and engagement? Static measures have previously been criticized for their lack of sensitivity to dynamic language processes (Campbell, Dollaghan, Needleman, & Janosky, 1997). The approach used in this study is to present children with a range of words from early- to late-acquired in the hopes of sampling a number of recently encountered new words for each child. The assumption is that words that have been encountered in the more distant past are further removed from the dynamic processes that supported learning and consequently fail to provide insights into those dynamic processes. In addition, these words are likely to be fully mastered with little variation in accuracy (i.e., highly correct). Likewise, words that have not yet been encountered fail to provide insights into word learning processes because they have not yet undergone word learning. These words are unknown with little variation in accuracy (i.e., highly incorrect). In contrast, recently encountered new words are those that have just been learned or are in the process of being learned. As a result, these words have the potential to provide insights into the dynamic processes that lead to their learning in the same way that recently exposed novel words in an experimental word learning paradigm provide insights into word learning processes. The difficulty when using a vocabulary probe is the tremendous variability in word learning experiences across children. That is, a word that is mastered for one child may be a recently learned word for a second child and a completely unknown word for a third child. For this reason, a wide range of words, in terms of both frequency of exposure and typical age-of-acquisition, needs to be tested to yield a sufficient sample of recently learned words for each child.

Purpose

The goal of the current study was to differentiate the effect of phonotactic probability on word learning from that of neighborhood density by fully crossing these two variables in stimuli selection. A second goal was to compare the independent and interactive effects of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

these two variables on word learning across preschool children differing in phonological status (i.e., children with phonological delays vs. children with typical development). A final goal was to accomplish these tasks using a more naturalistic and easily administered method, specifically an expressive and receptive vocabulary probe consisting of words sampling a range of frequencies and ages-of-acquisition.

Method

Participants

Fifty-four preschool children (age 3;5 - 6;7) participated: 20 with functional phonological delays (PD) and 34 with typical language development (TD). Based on parent report, none of the children had a history of cognitive, social, emotional, motor, visual, hearing, or major medical impairments. All children passed a hearing screening at study entry (ASHA, 1997). All children scored at or above the 16th percentile (1 standard deviation below the mean) on standardized tests of receptive and/or expressive vocabulary (Brownell, 2000a, 2000b). Test results for each group are displayed in Table 1.

The 20 children with PD (10 male; 10 female) met one of two possible criteria to be classified as having delayed phonological development. Children either scored at the 11th percentile or below on the *Goldman-Fristoe Test of Articulation - 2nd edition* (n = 17, Goldman & Fristoe, 2000) or scored between the 12th and 14th percentiles on the GFTA and had at least six target sounds with inventory or positional constraints based on analysis of productions on an extensive probe of English phonology (n = 3, Gierut, 2008). All 20 children evidenced normal development in language and cognition as defined by a score at the 16th percentile or above (1 standard deviation below the mean) on standardized tests of omnibus receptive language and nonverbal intelligence (Carrow-Woolfolk, 1995; Reynolds & Kamphaus, 2003).

The 34 children with TD (18 male; 16 female) evidenced typical phonological development as demonstrated by scores at or above the 24th percentile on the GFTA. As expected, t test comparisons showed that the groups differed significantly in their GFTA percentile ranks, t (52) = 9.05, $p < 0.001$. As shown in Table 1, the TD group was matched in gender, age, and raw receptive vocabulary scores to the PD group, χ^2 (54) = 0.04, $p > 0.80$ for gender and all t (52) < 0.90, all $p > 0.35$ for age and vocabulary.

Stimuli

Overview of Selection Procedures—There was a need to select words that represented a wide range of ages-of-acquisition so that recently learned words would be sampled for all children. Thus, the initial stimulus pool consisted of 442 real words compiled from previous age-of-acquisition studies (Carroll & White, 1973; Garlock, 1997; Snodgrass & Yuditsky, 1996). Phonotactic probability and neighborhood density were computed for each word and coded into four conditions using word length sensitive cut-points: (1) rare phonotactic probability - sparse neighborhood (n = 129); (2) rare phonotactic probability - dense neighborhood (n = 37); (3) common phonotactic probability - sparse neighborhood (n = 47); (4) common phonotactic probability - dense neighborhood (n = 86). Note that the number of potential stimuli varied across conditions. This is due to the previously documented correlation between phonotactic probability and neighborhood density (Storkel, 2004c; Vitevitch et al., 1999). Also, note that the number of stimuli is less than 442. This occurred for two reasons. First, two measures of phonotactic probability were used and the code for each measure had to agree (n = 115). Second, longer words had to be eliminated because there was no variation in density (n =28). Words were selected for each condition while matching variables related to vocabulary experience/exposure (i.e., age-of-acquisition, word

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

frequency), phonology (i.e., word length), and semantics (i.e., concreteness, familiarity, imagability) because it was hypothesized that these variables could influence responding, overshadowing effects of phonotactic probability and neighborhood density. These particular variables were chosen because of their ready availability for a large number of words. These procedures yielded 121 words (shown in Appendix A) across the four conditions: (1) rare phonotactic probability - sparse neighborhood (n = 35); (2) rare phonotactic probability - dense neighborhood (n = 24); (3) common phonotactic probability - sparse neighborhood (n = 27); (4) common phonotactic probability - dense neighborhood (n = 35).

After the data were collected using these 121 words, participant data were analyzed following the methods outlined in the results section. While numerous extraneous variables were controlled as previously described, some significant effects of phonotactic probability and neighborhood density in the participant analysis followed the trend for age-of-acquisition differences across conditions. Thus, the influence of age-of-acquisition could not be ruled-out. For this reason, 8 items were removed from the data set (i.e., those with an age-of-acquisition of 7 or greater) to achieve an even closer matching across conditions. Elimination of these items then required removal of 12 more items (i.e., those with length of 7 phonemes) to achieve a better match in word length across conditions. The 101 remaining items were distributed across the conditions in the following way: (1) rare phonotactic probability - sparse neighborhood (n = 27); (2) rare phonotactic probability - dense neighborhood (n = 23); (3) common phonotactic probability - sparse neighborhood (n = 20); (4) common phonotactic probability - dense neighborhood (n = 31). Removed items are marked in Appendix A. All following descriptive and inferential statistics refer only to this reduced set of items. Table 2 reports the descriptive statistics for all seven independent and control variables.

Independent Variables—Two measures of phonotactic probability were computed, positional segment sum and biphone sum, using an on-line calculator (Storkel, Hoover, & Kieweg, 2008). This on-line calculator computes positional segment sum, biphone sum, neighborhood density, and log word frequency based on a corpus of approximately 5,000 different words spoken by kindergarten or first grade children (Kolson, 1960; Moe, Hopkins, & Rush, 1982). In addition, the calculator provides the same calculations based on an adult corpus of approximately 20,000 different words from a dictionary (*Webster's Seventh Collegiate Dictionary*, 1967) and frequency in written language (Kucera & Francis, 1967). Note that calculations based on the child or adult corpus produced similar results, and only the child values are reported here. *Positional segment sum* is computed by adding the positional segment frequency for each phoneme in a word. Positional segment frequency is computed by adding the frequency of each word in the child corpus that contains a given phoneme in a given word position and then dividing by the sum of the frequency of every word in the dictionary that contains any phoneme in the same word position (Storkel, 2004c). *Biphone sum* is computed in a similar way but is based on pairs of adjacent sounds. Specifically, biphone sum is computed by adding the biphone frequency for each pair of phonemes in a word. Biphone frequency is computed by adding the frequency of each word in the child corpus that contains the given phoneme pair in the given word position and then dividing by the sum of the frequency of every word in the dictionary that contains any phoneme in the same word position (Storkel, 2004c).

To ensure that the rare and common conditions significantly differed on the measures of phonotactic probability while the sparse and dense conditions did not, positional segment sum and biphone sum were analyzing using a 2 Phonotactic Probability (rare, common) x 2 Neighborhood Density (sparse, dense) ANOVA. As shown in Table 2, common sound sequences had significantly higher positional segment sums than rare sound sequences, F (1,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

97) = 48.76, $p < 0.001$, $\eta_p^2 = 0.33$. Likewise, common sound sequences had significantly higher biphone sums than rare sound sequences, F (1, 97) = 47.68, $p < 0.001$, $\eta_p^2 = 0.33$. Also as intended, sparse and dense words had similar positional segment and biphone sums and the difference between rare and common sound sequences was similar across sparse and dense neighborhoods, all *Fs* (1, 97) < 3.20, all *ps* > 0.07, all η_p^2s < 0.04.

Neighborhood density is the number of words that differ from a given word by a one phoneme substitution, addition, or deletion (Storkel, 2004c). Neighborhood density was computed, using the previously described on-line child calculator (Storkel et al., 2008). To ensure that the sparse and dense conditions significantly differed in neighborhood density while the rare and common conditions did not, the number of neighbors was analyzed using a 2 Phonotactic Probability (rare, common) x 2 Neighborhood Density (sparse, dense) ANOVA. As shown in Table 2, dense neighborhoods had significantly more neighbors than sparse neighborhoods, F (1, 97) = 10.27, $p = 0.002$, $\eta_p^2 = 0.10$. Also as intended, rare and common sounds sequences had similar numbers of neighbors and the difference between sparse and dense neighborhoods was similar across rare and common sound sequences, all *Fs* (1, 97) < 0.60, all *ps* > 0.44, all η_p^2s < 0.01.

Control Variables—Age-of-acquisition (AoA) ratings were obtained from three sources (Carroll & White, 1973; Garlock, 1997; Snodgrass & Yuditsky, 1996). AoA ratings commonly are obtained by presenting words to adults and asking them to rate at what age they think they learned the word. The three AoA sources were selected because they were based on data from American English speakers using a similar 9-point rating scale for AoA. Across the three studies, a rating of 1 corresponded to an AoA of 0-2 years, a rating of 5 corresponded to an AoA of 6 years, and a rating of 9 corresponded to an AoA of 13+ years. Garlock (1997) and Snodgrass and Yuditsky (1996) included only these scale anchor points, whereas Carroll and White (1973) included additional anchor points (i.e., rating of 2 = AoA of 3 years, 3 = 4 years, 4 = 5 years, 6 = 7-8 years, 7 = 9-10 years, 8 = 11-12 years). If a given word occurred in more than one AoA source, the AoA across sources was averaged.

Log word frequency was obtained from the previously described on-line child calculator (Storkel et al., 2008). Frequency was taken from the original corpuses that were combined to create the calculator (Kolson, 1960; Moe et al., 1982). In the event that a word occurred in both corpuses, the raw frequencies from each corpus were added. The log base 10 was then computed and a constant value of 1 was added to each log frequency to avoid log frequencies of 0. Word frequency values were available for 81% of the selected probe words.

Word length was computed by counting the number of phonemes in the phonemic transcription provided by the previously described on-line child calculator (Storkel et al., 2008).

Semantic set size was obtained from an on-line database (Nelson, McEvoy, & Schreiber, 1998). Semantic set size was determined by presenting a printed word to a large group of adult participants and having each participant report the first word that came to mind that was meaningfully related to the given word. Responses reported by two or more participants are considered semantic neighbors of the word. The total number of different words reported as neighbors is the semantic set size. Semantic set size values were available for 79% of the selected probe words.

Concreteness ratings were obtained from an on-line database (Wilson, 1987). Concreteness values in this database were obtained from three sources (Gilhooly & Logie, 1980; Pavio, Yuille, & Madigan, 1968; Toglia & Battig, 1978). In general, concreteness ratings were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

obtained by asking adult participants to rate the concreteness of a given word on a 7-point scale where a high rating indicates "words referring to objects, materials, or persons" and a low rating indicates "words referring to abstract concepts that could not be experienced by the senses" (Gilhooly & Logie, 1980, p. 396). The on-line database converts the original ratings, multiplying them by 100 to avoid decimals. Concreteness values were available for 74% of the probe words.

Familiarity ratings were obtained from an on-line database (Wilson, 1987). Familiarity values in this database were obtained from three sources (Gilhooly & Logie, 1980; Pavio et al., 1968; Toglia & Battig, 1978). In general, familiarity ratings were obtained by asking adult participants to rate the familiarity of a given word on a 7-point scale where a rating of 7 indicates a words that is "seen, heard, or used every day" and a rating of 1 indicates a word that is "never seen, heard, or used" (Gilhooly & Logie, 1980, p. 396). The on-line database converts the original ratings, multiplying them by 100. Familiarity values were available for 80% of the probe words.

Imagability ratings were obtained from an on-line database (Wilson, 1987). Imagability values in this database were obtained from three sources (Gilhooly & Logie, 1980; Pavio et al., 1968; Toglia & Battig, 1978). In general, imagability ratings are obtained by asking adult participants to rate the imagability of a given word on a 7-point scale where a rating of 7 indicates "words arousing images most readily" and a rating of 1 indicates "words arousing images with great difficulty or not at all" (Gilhooly & Logie, 1980, p. 396). The on-line database converts the original ratings, multiplying them by 100. Imagability values were available for 74% of the probe words.

It was intended that there would be no difference in rare and common sound sequences or sparse and dense neighborhoods in age-of-acquisition, frequency, word length, semantic set size, concreteness, familiarity, or imagability. To examine this, each of the seven control variables was entered as the dependent variables in a 2 Phonotactic Probability (rare, common) x 2 Neighborhood Density (sparse, dense) ANOVA. Words with missing data were eliminated only from the ANOVA that required the missing data as the dependent variable. For all seven ANOVAs, there were no significant effects of phonotactic probability, all $Fs < 1.30$, all $ps > 0.25$, all η_p^2s < 0.02, or neighborhood density, all $Fs < 1.20$, all $ps > 0.25$, all η_p^2s < 0.02, or interactions of phonotactic probability and neighborhood density, all Fs < 1.90, all ps > 0.15, all η_p^2s < 0.03. Note that effect sizes for all control variables also were small (i.e., all η_p^2s < 0.03).

Differences in additional phonological variables, including word length in syllables, canonical structure, and age of consonant acquisition (Smit, 1993; Smit, Hand, Freilinger, Bernthal, & Bird, 1990), were examined in a similar manner (see Appendix B) with no significant effects of phonotactic probability -- neighborhood density conditions, all $Fs < 2.70$ and all χ^2s < 5.10, all ps > 0.10, all η_p^2 s < 0.03. Note that effect sizes for these additional phonological variables also were small (i.e., all η_p^2s < 0.03).

Probe construction and administration

For each of the selected real words, two color pictures were obtained from a variety of books and on-line resources. One picture was randomly assigned to the expressive probe and one picture was randomly assigned to the receptive probe. The expressive probe was always administered before the receptive probe because it was thought that hearing the name of stimuli in the receptive probe could influence responding on the expressive probe.

For the expressive probe, the selected pictures were randomized and inserted into a PowerPoint file for presentation to participants. Each picture was presented individually and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

participants were prompted to name the picture. Responses were scored as correct if the participant produced a recognizable attempt at the target word (i.e., exactly correct articulation was not required). Accurate articulation was not required because this might unfairly penalize the children with phonological delays who were expected to make more articulation errors than the typically developing children. In other words, the goal was to test expressive vocabulary, not articulation. Phonological analyses were consulted to determine children's typical error patterns, and this information assisted in scoring. In addition, words that were semantically similar to the target word (e.g., synonyms, superordinate categories) were scored as incorrect because these types of errors would likely differ from the target on the manipulated variables (i.e., phonotactic probability and neighborhood density). Thus, *accuracy* on the expressive probe refers to lexical accuracy rather than phonological or semantic accuracy.

For the receptive probe, the selected pictures were randomized and inserted into a PowerPoint file for presentation to participants. In addition, three foils were selected for each target. One foil was a picture of a semantically related item, which was defined as another item from the same superordinate category (but differing in sound structure). The second foil was a picture of a phonologically related item, which was defined as a real word that shared the same initial phoneme (but differed in superordinate category). The third foil was an unrelated picture that did not share superordinate category or initial phoneme or rhyme with the target word. Placement of targets and foils on a given PowerPoint slide was randomized across items. Each set of four pictures was presented individually. The examiner asked the child to point to the picture that corresponded to the target word. The participant pointed to one of the pictures, and the choice was scored by the examiner.

Results

Correlations between demographic variables (i.e., chronological age, raw receptive vocabulary score, and raw expressive vocabulary score) and experimental variables (i.e., proportion correct for each phonotactic probability - neighborhood density condition for each type of probe) were examined to determine whether covariates needed to be used to address the experimental questions. Recall that past research has shown that the effects of phonotactic probability and neighborhood density may be modulated by age or vocabulary. Although the PD and TD groups were matched on age and vocabulary, both groups exhibited a wide range of ages and vocabulary scores (see Table 1). As shown in Table 3, chronological age, raw receptive vocabulary score, and raw expressive vocabulary score were significantly positively correlated with proportion correct in all experimental conditions. Specifically, proportion correct on the experimental tasks tended to increase as chronological age, raw receptive vocabulary scores, or raw expressive vocabulary scores increased.

A series of partial correlation analyses was then conducted to determine whether controlling for one demographic variable could reduce the correlation between the remaining demographic variables and the experimental variables to yield an optimal covariate. After partialing out effects of chronological age, 12 of 16 possible correlations between demographic (i.e., raw receptive vocabulary score, raw expressive vocabulary score) and experimental variables remained significant. After partialing out effects of raw receptive vocabulary scores, none of the correlations between demographic (i.e., chronological age, raw expressive vocabulary score) and experimental variables remained significant, all rs < 0.25, all ps > 0.07, all r^2s < 0.07. After partialing out effects of raw expressive vocabulary score, 10 of 16 possible correlations between demographic (i.e., chronological age, raw receptive vocabulary score) and experimental variables remained significant. Thus, it was determined that raw receptive vocabulary score was the optimal covariate for all remaining

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

analyses. Raw receptive vocabulary scores were mean centered, as is typical in ANCOVA. The main effect of the covariate was significant in all of the following ANCOVA analyses, all *F*s > 22.25, all *p*s < 0.001, all η_p^2s > 0.40.

Proportion correct on the vocabulary probe was analyzed using a 2 Phonotactic Probability (rare, common) x 2 Neighborhood Density (sparse, dense) x 2 Probe Type (expressive, receptive) x 2 Group (PD, TD) ANCOVA with mean centered raw receptive vocabulary score as the covariate. *P*-critical was set at 0.025 because this produced consistent effects across multiple analyses systematically removing specific items (see stimuli section of methods). Results showed a significant effect of probe type, F (1, 51) = 876.68, $p < 0.001$, $\eta_p^2 = 0.95$, with responses to the expressive probe ($M = 0.60$, $SD = 0.11$, range = 0.30-0.90) being less accurate than responses to the receptive probe ($M = 0.87$, $SD = 0.08$, range = 0.56-1.00). This is expected given the difference in response format with the expressive probe having an open-response format (i.e., participant must recall the correct answer) and the receptive probe having a closed-response format (i.e., participant must recognize/select the correct answer). This leads to the expressive probe being more difficult than the receptive probe (Clopper, Pisoni, & Tierney, 2006). This main effect of probe type is observed in all remaining analyses, all *F*s > 235.45, all *p*s < 0.001, all η_p^2s > 0.90, but will not be specifically reported or commented on further. In addition, probe type and the covariate receptive vocabulary showed a significant interaction in this analysis, F (1, 51) = 15.00, $p < 0.001$, $\eta_p^2 = 0.23$, and in all remaining analyses, all *F*s > 5.70, all *p*s < 0.025, all η_p^2s > 0.10. In all cases, the difference between performance on the expressive and receptive probes decreased as scores on the covariate receptive vocabulary test increased, $r = -0.20$ - -0.53, $r^2 = 0.04$ - 0.28. Again, this effect will not be specifically reported in the remaining analyses because it is not the main focus of the research.

Turning to effects related to the main research questions, several interactions involving phonotactic probability and neighborhood density were obtained, including (1) phonotactic probability and neighborhood density, F (1, 51) = 9.07, $p < 0.01$, $\eta_p^2 = 0.15$; (2) phonotactic probability, neighborhood density, and the covariate receptive vocabulary, F (1, 51) = 6.21, $p < 0.05$, $\eta_p^2 = 0.11$; (3) phonotactic probability and probe type, F (1, 51) = 20.70, $p < 0.001$, $\eta_p^2 = 0.29$.

Four ANCOVAs were conducted to unpack these significant interactions. The first set of two ANCOVAs examined the effect of phonotactic probability within each level of neighborhood density (sparse, dense) using a 2 Phonotactic Probability (rare, common) x 2 Probe Type (expressive, receptive) x 2 Group (PD, TD) ANCOVA with mean centered raw receptive vocabulary score as the covariate. The second set of two ANCOVAs examined the effect of neighborhood density within each level of phonotactic probability (rare, common) using a 2 Neighborhood Density (sparse, dense) x 2 Probe Type (expressive, receptive) x 2 Group (PD, TD) ANCOVA with mean centered raw receptive vocabulary score as the covariate. As previously noted, *p*-critical was set at 0.025 because this tended to produce consistent effects across multiple analyses systematically removing specific items.

Effect of Phonotactic Probability

Sparse neighborhoods—For sparse neighborhoods, responses to rare sound sequences ($M = 0.75$, $SD = 0.16$, range = 0.41-1.00) were significantly more accurate than responses to common sound sequences ($M = 0.73$, $SD = 0.19$, range = 0.35-1.00), F (1, 51) = 5.61, $p < 0.025$, $\eta_p^2 = 0.10$. This main effect was qualified by a significant interaction with probe type, F (1, 51) = 8.40, $p < 0.01$, $\eta_p^2 = 0.14$. As shown in Table 4, responses to rare sound sequences ($M = 0.63$, $SD = 0.10$, range = 0.41-0.85) were more accurate than responses to common sound sequences ($M = 0.58$, $SD = 0.12$, range = 0.35-0.90) in the expressive probe, F (1, 51) = 12.49, $p = 0.001$, $\eta_p^2 = 0.20$. In contrast, responses to rare sound sequences (M =

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

0.88, *SD* = 0.08, range = 0.56-1.00) and to common sound sequences (*M* = 0.88, *SD* = 0.09, range = 0.65-1.00) were similarly accurate in the receptive probe, F (1, 51) = 0.03, $p > 0.85$, $\eta_p^2 < 0.01$. Taken together, in sparse neighborhoods, children knew more words composed of rare sound sequences than words composed of common sound sequences, but only on the expressive probe.

Dense neighborhoods—For dense neighborhoods, there was no main effect of phonotactic probability, F (1, 51) = 2.44, $p > 0.10$, $\eta_p^2 < 0.05$. However, phonotactic probability showed significant interactions with group, F (1, 51) = 6.14, $p < 0.025$, $\eta_p^2 = 0.11$, probe type, F (1, 51) = 8.77, $p < 0.01$, $\eta_p^2 = 0.15$, and the receptive vocabulary covariate, F (1, 51) = 5.54, $p < 0.025$, $\eta_p^2 = 0.10$. Follow-up analyses were conducted for each group (PD vs. TD). As shown in Table 4 for the PD group, responses to rare sound sequences (*M* = 0.73, *SD* = 0.17, range = 0.30-1.00) and to common sound sequences (*M* = 0.73, *SD* = 0.18, range = 0.39-1.00) were similarly accurate, F (1, 18) = 0.34, $p > 0.55$, $\eta_p^2 < 0.02$. In contrast, there was a significant effect of phonotactic probability for the TD group, F (1, 32) = 10.88, $p < 0.01$, $\eta_p^2 = 0.25$, but this was qualified by a significant interaction with probe type, F (1, 32) = 7.30, $p < 0.025$, $\eta_p^2 = 0.19$. As shown in Table 4, there was no significant effect of phonotactic probability for the TD group in the expressive probe, F (1, 32) = 0.21, $p > 0.60$, $\eta_p^2 < 0.01$, whereas the TD group responded to common sound sequences (*M* = 0.89, *SD* = 0.06, range = 0.77-1.00) more accurately than rare sound sequences (*M* = 0.83, *SD* = 0.07, range = 0.70-0.96) in the receptive probe, F (1, 32) = 46.05, $p < 0.001$, $\eta_p^2 = 0.59$. In summary, for dense neighborhoods, only children with typical development knew more words composed of common sound sequences than words composed of rare sound sequences, and this was evident only on the receptive probe.

Effect of Neighborhood Density

Rare sound sequences—For rare sound sequences as shown in Table 4, responses to sparse neighborhoods (*M* = 0.75, *SD* = 0.16, range = 0.41-1.00) were more accurate than responses to dense neighborhoods (*M* = 0.72, *SD* = 0.16, range = 0.30-1.00), F (1, 51) = 13.15, $p = 0.001$, $\eta_p^2 = 0.21$.

Common sound sequences—For common sound sequences as shown in Table 4, responses to sparse neighborhoods (*M* = 0.73, *SD* = 0.19, range = 0.35-1.00) and to dense neighborhoods (*M* = 0.74, *SD* = 0.17, range = 0.39-1.00) were similarly accurate, F (1, 51) = 1.09, $p > 0.25$, $\eta_p^2 < 0.03$.

Discussion

The goals of this study were to differentiate the effects of phonotactic probability and neighborhood density on a naturalistic probe of word learning administered to two groups of preschool children differing in phonological development (i.e., children with phonological delays vs. children with typical development) but matched on age and receptive vocabulary scores. Results showed that the effect of phonotactic probability was dependent on neighborhood density (sparse vs. dense), probe type (expressive vs. receptive), and phonological status (phonological delay vs. typical development). In contrast, the effect of neighborhood density was dependent on phonotactic probability (rare vs. common) alone. In general, similar effects of phonotactic probability and neighborhood density were observed across children differing in phonological development, with the exception of the effect of phonotactic probability in dense neighborhoods. Taken together, the results suggest that more traditional vocabulary probes may be sensitive to the role of phonotactic probability and neighborhood density in word learning. Each of these three issues will be considered in turn.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Role of phonotactic probability and neighborhood density in typical development

Results suggest variability in the role of phonotactic probability in word learning by typically developing children with the direction of the effect of phonotactic probability depending on the neighborhood density of the word to be learned and the type of probe. Specifically, typically developing children learned rare sound sequences more readily than common sound sequences when the neighborhood was sparse but only on the expressive probe. The reverse pattern, with common sound sequences being learned more readily than rare sound sequences, was observed when the neighborhood was dense but only on the receptive probe. Likewise, the role of neighborhood density in word learning depended on the phonotactic probability of the words to be learned. Specifically, typically developing children learned words in sparse neighborhoods more readily than words in dense neighborhoods but only for rare phonotactic probability. No effect of neighborhood density was observed for common phonotactic probability. An account of the phonotactic probability and neighborhood density effects will be presented first, followed by an account of the task differences.

Taken together, phonotactic probability and neighborhood density effects converged such that the optimal conditions were rare phonotactic probability with sparse neighborhoods and common phonotactic probability with dense neighborhoods. This seems like an apparent contradiction with words that are more distinctive (i.e., rare phonotactic probability with sparse neighborhoods) being learned readily and words that are more typical (i.e., common phonotactic probability with dense neighborhoods) being learned readily. How is it that distinctive and typical words can both facilitate word learning? One possibility is that these endpoints of the continuum affect different hypothesized components of word learning. Specifically, distinctive words may more efficiently trigger word learning (Storkel et al., 2006). That is, a rare sound sequence will activate existing phonological representations but these activated phonological representations will activate few existing lexical representations because rare sound sequences, by definition, occur infrequently in the language. Likewise, a new word in a sparse neighborhood will activate few existing lexical representations. Because of this minimal lexical activation, determining that none of the existing lexical representations exactly matches the novel word will likely occur rapidly and accurately, efficiently triggering learning of the new word. This hypothesis warrants direct testing using a task that can unambiguously tap triggering, such as any novelty detection task (e.g., Merriman & Marazita, 1995).

In contrast, more typical words may facilitate configuration, namely the creation of a new representation in the lexicon (Leach & Samuel, 2007). Common sound sequences are easier to hold in working memory (Gathercole et al., 1999; Thorn & Frankish, 2005). Likewise, words from dense neighborhoods are easier to hold in working memory (Roodenrys & Hinton, 2002; Thomson et al., 2005; Thorn & Frankish, 2005). In this case, a more complete and accurate representation of the sound sequence will be held in working memory for novel words that are both common and dense, supporting creation of a more complete and accurate lexical representation in long term memory for these words.

In addition, more typical words may facilitate engagement, specifically the integration of new representations with existing representations (Leach & Samuel, 2007). In terms of phonotactic probability, common sound sequences will activate existing phonological representations which will spread activation to many lexical representations, including that of the new word. These lexical representations will spread activation back to existing phonological representations. This interactive process serves to strengthen the connections between phonological and lexical representations, with this strengthening being greater for common sound sequences than rare sound sequences. In terms of neighborhood density, integration of a new representation with many existing representations, as would occur in a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

dense neighborhood, could strengthen the new lexical representation (Storkel et al., 2006). These hypotheses concerning the effect of neighborhood density on configuration and engagement warrant direct testing, using methods that can unambiguously disentangle configuration (e.g., forced-choice recognition tasks, threshold discrimination tasks) and engagement (e.g., lexical decision, pause detection, see Gaskell & Dumay, 2003; Leach & Samuel, 2007).

Taken together, rare sound sequences and sparse neighborhoods (i.e., distinctive words) may have provided converging cues to facilitate triggering of word learning, whereas common sound sequences and dense neighborhoods (i.e., typical words) may have provided converging cues to facilitate configuration and/or engagement. This is somewhat consistent with the findings in adult word learning from Storkel and colleagues (2006). However, adults appeared to have a clearer division of labor between phonotactic probability and neighborhood density than the typically developing children in the current study, showing no significant interaction between the two variables. That is, for adults rare sound sequences facilitated triggering regardless of neighborhood density, and dense neighborhoods facilitated configuration and/or engagement regardless of phonotactic probability. In contrast, the children in the current study appeared to benefit from a convergence of phonotactic probability and neighborhood density for triggering, configuration, and engagement. This suggests that a critical part of development in word learning may be a re-weighting of cues for triggering, configuration, and engagement such that a smaller set of cues is used more heavily for a given component of word learning.

Another interesting note about the effect of phonotactic probability is that expressive and receptive vocabulary probes differed in their sensitivity to phonotactic probability. It is unclear whether differences across probe type should be interpreted as revealing important underlying word learning processes or as resulting from methodological differences. Considering first the hypothesis that task differences reveal something about the word learning process, one must consider how the expressive and receptive probes differentially tap underlying representations. On the expressive probe, children are shown a picture. This picture presumably activates a semantic representation in long-term memory, which in turn activates lexical and phonological representations. These lexical and phonological representations must be relatively accurate and detailed to support a recognizable attempt in producing the target word. In contrast, on the receptive probe, children are shown multiple pictures and hear the target word. Hearing the target word presumably activates phonological and lexical representations, which in turn activate a semantic representation. The semantic representation in long-term memory must then be compared to the picture choices so that a matching picture can be selected. Note that the receptive probe does not require complete, accurate, and detailed representations to support a correct response. Existing lexical or semantic representations need only have enough accurate information to support retrieval of a single unique lexical representation (i.e., only one lexical representation completely or partially matches the spoken word) and a single semantic representation that matches one of the picture choices. This hypothesized difference in the level of detail needed in representations to support a correct response across tasks is consistent with the obtained main effect of task (i.e., accuracy on the expressive task was always worse than accuracy on the receptive task). This hypothesis also suggests that the expressive task may more directly tap the quality or level of detail in lexical and semantic representations than the receptive task, whereas the receptive task may more directly tap the association between a lexical and semantic representation. In this way, the expressive task may be more sensitive to triggering, whereas the receptive task may be more sensitive to engagement, which includes the formation of associations between different representations, such as lexical and semantic (Leach & Samuel, 2007).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Turning to potential methodological differences, although the same words were used on both the expressive and receptive vocabulary probes, the pictures differed across the probes. No attempt was made to examine the equivalence of pictures across probes, other than to have unfamiliar adults attempt to identify the pictures to ensure that the pictures clearly depicted the target word. However, it is possible that unmeasured differences across pictures influenced responding. Given these concerns, the previous theoretical interpretation of differences in the effect of phonotactic probability across expressive and receptive probes should be viewed with caution. Replication clearly is warranted.

Comparison between delayed versus typical development

Children with phonological delays showed similar effects of rare phonotactic probability and sparse neighborhoods as children with typical development. Thus, children with phonological delays appear to benefit from the converging cues of rare phonotactic probability and sparse neighborhoods to trigger word learning in a manner similar to typically developing children. Moreover, it seems that the triggering component of word learning may be relatively intact in children with phonological delays. However, children with phonological delays did not show the same benefit of common phonotactic probability and dense neighborhoods as children with typical development. This is consistent with the findings of Storkel (2004b), where children with phonological delays performed more poorly on common dense sound sequences than rare sparse sequences. Moreover, this finding suggests that children with phonological delays may differ from children with typical development in the configuration and/or engagement components of word learning. In terms of configuration, it is possible that phonotactic probability and neighborhood density do not affect working memory in children with phonological delays in the same manner as in children with typical development. This seems somewhat unlikely given that children with phonological delays do show better performance for common sound sequences than for rare sound sequences in nonword repetition tasks (Munson, Edwards, & Beckman, 2005b); however, it is possible that differences could arise if phonotactic probability and neighborhood density were fully crossed in a working memory task. For this reason, differences in configuration can not be ruled out. In terms of engagement, it is possible that high levels of interactive activation between phonological and lexical representations as well as integration of new representations with many existing representations do not benefit children with phonological delays. Presumably, these children have weaker phonological representations and may have less detailed lexical representations (c.f., Edwards, Fourakis, Beckman, & Fox, 1999; Edwards, Fox, & Roger, 2002). For this reason, high levels of activation may overwhelm the system, leading to confusion between new and existing representations, thereby reducing the typical benefits of common sound sequences and dense neighborhoods.

Sensitivity of traditional vocabulary tests

These results suggest that traditional vocabulary tests can be used to examine the role of phonotactic probability and neighborhood density in word learning. Moreover, the findings indicate that traditional vocabulary tests may be sensitive to the components of the word learning process, specifically triggering, configuration, and engagement. This is an important issue because past work has suggested that traditional vocabulary tests, with their emphasis on the products of word learning, may not be sensitive to the word learning process itself (Campbell et al., 1997). However, the findings from the current vocabulary probe match other studies that more directly test word learning processes (Storkel, 2001, 2003, 2004b; Storkel et al., 2006; Storkel & Maekawa, 2005), suggesting that at least under certain circumstances word learning processes may be revealed by traditional vocabulary probes. In particular, probe construction likely is critical. For the probe in the current study, items from each phonotactic probability/neighborhood density condition were selected to be

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

matched on age-of-acquisition and a range of ages-of-acquisition was used. The importance of this control of age-of-acquisition was that for a given child there would likely be three sets of words administered: (1) early learned words that were highly accurate; (2) recently learned words that varied in accuracy; (3) yet to be learned words that were highly inaccurate. It is the recently learned words that have the same potential to reveal word learning processes as nonword learning paradigms used in other experiments because these words are closer to the dynamic processes involved in their learning. This hypothesis warrants further study but suggests that static measures of vocabulary do have the potential to reveal fine grain information about word learning.

Conclusions

This study demonstrates that effects of phonotactic probability and neighborhood density can be detected using traditional clinical methods for assessing vocabulary. Moreover, results showed that typically developing children require a convergence of phonotactic probability and neighborhood density to support word learning and that the necessary convergence may vary across components of word learning. Specifically, rare sound sequences and sparse neighborhoods were hypothesized to facilitate triggering of learning, whereas common sound sequences and dense neighborhoods were hypothesized to facilitate configuration and engagement. Children with phonological delays showed similar patterns related to triggering of learning, but demonstrated potential differences in configuration and/or engagement. In particular, children with phonological delays did not appear to benefit from common sound sequences and dense neighborhoods in the same way as typically developing children. Results further suggest that carefully constructed vocabulary probes may be sensitive to the word learning process.

Acknowledgments

This research was supported by NIH Grants DC06545, DC08095, DC00052, DC009135, DC05803, and HD02528. The following individuals contributed to stimulus creation, data collection, data processing, and reliability calculations: Teresa Brown, Jennie Fox, Andrea Giles, Stephanie Gonzales, Nicole Hayes, Shannon Rogers, Josie Row, Katie Shatzer, Maki Sueto, Courtney Winn, and Emily Zimmerman.

Appendix A: Vocabulary Probe Words

Rare Phonotactic Probability		**Common Phonotactic Probability**	
Sparse n = 35	**Dense n = 24**	**Sparse n = 27**	**Dense n = 35**
bagpipe	anchor	ant	basket
beaver	ball	banjo	bear
chef	bird	cactus	belt
chisel[a]	boot	canteen	blender
clogs[a]	broom	carrot	bread
cloud	duck	cow	bullet
clown	feather	desk	bus
couch	flute	doll	camel
dog	leaf	dress	candle
donkey	leopard	fence	car
fish	lock	flask[a]	caster[a]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Rare Phonotactic Probability		Common Phonotactic Probability	
Sparse n = 35	**Dense n = 24**	**Sparse n = 27**	**Dense n = 35**
flashlight[b]	nail	hammer	deer
frog	needle	hanger	elephant[b]
glass	peach	hydrant[b]	fan
globe	rope	jet	hair
guitar	skunk	lemon	harp
knife	spool	lettuce	hill
leg	squirrel	pencil	ladder
light switch[b]	swing	penguin[b]	lobster
monkey	table	pepper	mitten
motel	thimble	pig	mountain
mouse	turtle	poncho[a]	necklace
mushroom	waterbed[a]	sandwich[b]	nun
peacock	whistle	spoon	pants
sheep		tree	parrot
shirt		trumpet[b]	pear
shoe		windmill[b]	pen
surfboard[a]			propeller[b]
syringe[a]			pumpkin[b]
tiger			sun
tights			tent
toothbrush[b]			toaster
vase			toe
watch			trunk
wineglass[b]			vest

[a]Items removed from analysis due to late AoA.

[b]Items removed from analysis due to length.

Appendix B: Additional Phonological Characteristics of Probe Words Included in the Analysis

	Rare Phonotactic Probability		Common Phonotactic Probability	
	Sparse	**Dense**	**Sparse**	**Dense**
Word Length: Proportion of Words by Number of Syllables				
1-Syllable	67%	61%	50%	61%
2-Syllable	33%	39%	50%	39%
Canonical Structure: Proportion of Words for the Most Frequent 1-Syllable Structures				
CVC	61%	64%	30%	58%

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

	Rare Phonotactic Probability		Common Phonotactic Probability	
	Sparse	**Dense**	**Sparse**	**Dense**
CCVC	28%	29%	20%	5%
CVCC	6%	0%	20%	26%
Canonical Structure: Proportion of Words for the Most Frequent 2-Syllable Structures				
CVCV	22%	44%	30%	25%
CVCVC	33%	11%	30%	17%
CVCCV	22%	11%	20%	25%
Canonical Structure: Proportion of Words with a Cluster by Word Position				
Word Initial	19%	30%	15%	10%
Word Final	4%	4%	15%	19%
Age of Consonant Acquisition in Years: Means (and Standard Deviations) by Word Position				
Word Initial	4.4 (1.6)	5.0 (1.9)	4.3 (1.8)	3.9 (1.5)
Word Final	4.9 (1.7)	4.9 (1.9)	4.9 (1.8)	5.5 (2.2)

Statistical analysis of all of the above variables failed to detect significant differences across conditions, all *F*s < 2.70 and all χ^2s < 5.10, all *p*s > 0.10, all η_p^2 s < 0.03.

References

ASHA. Guidelines for screening for hearing impairment-preschool children, 3-5 years. Asha 1997;4:IV-74cc–IV-74ee.

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

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

Campbell T, Dollaghan C, Needleman H, Janosky J. Reducing bias in language assessment: Processing-dependent measures. Journal of Speech, Language, and Hearing Research 1997;40:519–525.

Carroll JB, White MN. Age-of-acquisition norms for 220 picturable nouns. Journal of Verbal Learning and Verbal Behavior 1973;12:563–576.

Carrow-Woolfolk, E. Oral and Written Language Scales. American Guidance Service, Inc.; Circle Pines, MN: 1995.

Clopper CG, Pisoni DB, Tierney AT. Effects of open-set and closed-set task demands on spoken word recognition. Journal of the American Academy of Audiology 2006;17(5):331–349. [PubMed: 16796300]

Coady JA, Aslin RN. Phonological neighbourhoods in the developing lexicon. Journal of Child Language 2003;30(2):441–470. [PubMed: 12846305]

Dell GS. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language 1988;27(2):124–142.

Edwards J, Beckman ME, Munson B. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. Journal of Speech, Language and Hearing Research 2004;47:421–436.

Edwards J, Fourakis M, Beckman ME, Fox RA. Characterizing knowledge deficits in phonological disorders. Journal of Speech, Language, and Hearing Research 1999;42:169–186.

Edwards J, Fox R, Roger C. Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy. Journal of Speech and Hearing Research 2002;45:231–242.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Felsenfeld S, Broen PA, McGue M. A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. Journal of Speech and Hearing Research 1992;35:1114–1125. [PubMed: 1280310]

Garlock, VM. Unpublished Doctoral dissertation. University of Alabama; Birmingham: 1997. The development of spoken word recognition and phoneme awareness during the preliteracy and early literacy periods: A test of the lexical restructuring model.

Garlock VM, Walley AC, Metsala JL. Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. Journal of Memory and Language 2001;45(3):468–492.

Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. Cognition 2003;89(2):105–132. [PubMed: 12915296]

Gathercole SE, Frankish CR, Pickering SJ, Peaker S. Phonotactic influences on short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition 1999;25:84–95.

Gierut, JA. Phonological disorders and the developmental phonology archive. In: Dinnsen, DA.; Gierut, JA., editors. Optimality theory, phonological acquisition and disorders. Equinox; London: 2008. p. 37-92.

Gilhooly KJ, Logie RH. Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. Behavior Research Methods, Instruments and Computers 1980;12:395–427.

Goldman, R.; Fristoe, M. Goldman-Fristoe Test of Articulation-2. American Guidance Service; Circles Pines, MN: 2000.

Gupta P, MacWhinney B. Vocabulary acquisition and verbal short-term memory: Computational and neural bases. Brain and Language. Special Issue: Computer models of impaired language 1997;59(2):267–333.

Kolson, CJ. Unpublished Doctoral Dissertation. University of Pittsburgh; Pittsburgh: 1960. The vocabulary of kindergarten children.

Kucera, H.; Francis, WN. Computational analysis of present-day American English. Brown University; Providence, RI: 1967.

Leach L, Samuel AG. Lexical configuration and lexical engagement: When adults learn new words. Cognitive Psychology 2007;55:306–353. [PubMed: 17367775]

Levelt, WJM. Speaking: From intention to articulation. The MIT Press; Cambridge, MA, US: 1989. p. xivp. 566ACL-MIT Press series in natural-language processing1989

Luce PA, Goldinger SD, Auer ET Jr. Vitevitch MS. Phonetic priming, neighborhood activation, and PARSYN. Perception & Psychophysics 2000;62(3):615–625. [PubMed: 10909252]

Maekawa J, Storkel HL. Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. Journal of Child Language 2006;33:439–459. [PubMed: 17017275]

Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: Studies with artificial lexicons. Journal of Experimental Psychology: General 2003;132(2):202–227. [PubMed: 12825637]

Mainela-Arnold E, Evans JL, Coady JA. Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. Journal of Speech and Hearing Research 2008;51:381–393.

McClelland JL, Elman JL. The TRACE model of speech perception. Cognitive Psychology 1986;18(1):1–86. [PubMed: 3753912]

Merriman WE, Marazita JM. The effect of hearing similar-sounding words on young 2-year-olds' disambiguation of novel noun reference. Developmental Psychology 1995;31(6):973–984.

Metsala JL. An examination of word frequency and neighborhood density in the development of spoken-word recognition. Memory and Cognition 1997;25:47–56.

Moe, AJ.; Hopkins, KJ.; Rush, RT. The vocabulary of first grade children. Thomas; Springfield, IL: 1982.

Munson B, Edwards J, Beckman ME. Phonological knowledge in typical and atypical speech-sound development. Topics in Language Disorders 2005a;25(3):190–206. [PubMed: 20209070]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Munson B, Edwards J, Beckman ME. Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. Journal of Speech, Language, and Hearing Research 2005b;48(1):61–78.

Munson B, Swenson CL, Manthei SC. Lexical and phonological organization in children: Evidence from repetition tasks. Journal of Speech, Language and Hearing Research 2005;48(1):108–124.

Nelson, DL.; McEvoy, C.; Schreiber, T. The University of South Florida word association, rhyme, and word fragment norms. 1998. Retrieved Feb 1, 2005, from http://www.usf.edu/FreeAssociation/

Newman RS, German DJ. Life span effects of lexical factors on oral naming. Language and Speech 2005;48(2):123–156. [PubMed: 16411502]

Norris D. Shortlist: A connectionist model of continuous speech recognition. Cognition 1994;52(3): 189–234.

Pavio A, Yuille JC, Madigan SA. Concreteness, imagery and meaningfulness values for 925 words. Journal of Experimental Psychology Monograph Supplement 1968;76(3, part 2)

Reynolds, CR.; Kamphaus, RW. Reynolds Intellectual Assessment Scales and the Raynolds Intellectual Screening Test. Psychological Assessment Resources, Inc.; Lutz, FL: 2003.

Roodenrys S, Hinton M. Sublexical or lexical effects on serial recall of nonwords? Journal of Experimental Psychology: Learning, Memory, and Cognition 2002;28(1):29–33.

Shriberg LD, Kwiatkowski J. Developmental phonological disorders. I: A clinical profile. Journal of Speech and Hearing Research 1994;37(5):1100–1126. [PubMed: 7823556]

Shriberg LD, Kwiatkowski J, Best S, Hengst J, Terselic-Weber B. Characteristics of children with phonological disorders of unknown origin. Journal of Speech and Hearing Disorders 1986;51:140–161. [PubMed: 3702362]

Smit AB. Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: Word-initial consonant clusters. Journal of Speech and Hearing Research 1993;36:931–947. [PubMed: 8246482]

Smit AB, Hand L, Freilinger JJ, Bernthal JE, Bird A. The Iowa Articulation Norms Project and its Nebraska replication. Journal of Speech and Hearing Disorders 1990;55(4):779–798. [PubMed: 2232757]

Snodgrass JG, Yuditsky T. Naming times for the Snodgrass and Vanderwart pictures. Behavior Research Methods, Instruments and Computers 1996;28:516–536.

Storkel HL. Learning new words: Phonotactic probability in language development. Journal of Speech, Language, and Hearing Research 2001;44(6):1321–1337.

Storkel HL. Learning new words II: Phonotactic probability in verb learning. Journal of Speech, Language, and Hearing Research 2003;46(6):1312–1323.

Storkel HL. Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. Applied Psycholinguistics 2004a;25(2):201–221.

Storkel HL. The emerging lexicon of children with phonological delays: Phonotactic constraints and probability in acquisition. Journal of Speech, Language, and Hearing Research 2004b;47(5):1194–1212.

Storkel HL. 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 2004c;47(6):1454–1468.

Storkel HL. Developmental differences in the effects of phonological, lexical, and semantic variables on word learning by infants. Journal of Child Language 2009;36:291–321. [PubMed: 18761757]

Storkel HL, Armbruster J, Hogan TP. Differentiating phonotactic probability and neighborhood density in adult word learning. Journal of Speech, Language, and Hearing Research 2006;49(6): 1175–1192.

Storkel, HL.; Hoover, JR.; Kieweg, D. The child mental lexicon: Neighborhood density, phonotactic probability, and word frequency online calculator. 2008. Publication. Retrieved 3/30/2009: http://www.bncdnet.ku.edu/cgi-bin/DEEC/info_ccc.vi

Storkel HL, Maekawa J. A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. Journal of Child Language 2005;32(4):827–853. [PubMed: 16429713]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Thomson JM, Richardson U, Goswami U. Phonological similarity neighborhoods and children's short-term memory: typical development and dyslexia. Memory & Cognition 2005;33(7):1210–1219.

Thorn AS, Frankish CR. Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. Journal of Experimental Psychology: Learning, Memory, and Cognition 2005;31(4):729–735.

Toglia, MP.; Battig, WR. Handbook of semantic word norms. Erlbaum; New York: 1978.

Vitevitch MS. The influence of phonological similarity neighborhoods on speech production. Journal of Experimental Psychology: Learning, Memory, and Cognition 2002;28:735–747.

Vitevitch MS, Armbruster J, Chu S. Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. Journal of Experimental Psychology: Learning, Memory, and Cognition 2004;30(2):1–16.

Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory of Language 1999;40:374–408.

Vitevitch MS, Luce PA, Pisoni DB, Auer ET. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain and Language. Special Issue: Mental lexicon 1999;68(1-2):306–311.

Webster's Seventh Collegiate Dictionary. Library Reproduction Service; Los Angeles: 1967.

Wilson, MD. MRC Psycholinguistic Database: Machine Usable Dictionary. Version 2.00. 1987. Publication. Retrieved 2/1/2005: http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm

Zamuner TS, Gerken L, Hammond M. Phonotactic probabilities in youn children's speech production. Journal of Child Language 2004;31:515–536. [PubMed: 15612388]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Participant characteristics

	Children with phonological delays (n = 20)	Children with typical development (n = 34)
Gender	50% male 50% female	53% male 47% female
Age *M* *SD* range	4;9 0;10 3;5 - 6;7	4;7 0;8 3;6 - 6;4
GFTA percentile**	6 4 1 - 14	57 25 24-98
ROWPVT raw score	58 14 33-82	58 9 42-76
ROWPVT standard score	104 10 85 - 120	105 7 90 - 123
EOWPVT standard score	103 10 86 - 117	104 8 83 - 121
OWLS receptive standard score	99 10 85 - 116	N/A
RIST standard score	115 20 89 - 155	N/A

Note. GFTA = Goldman-Fristoe Test of Articulation - 2, ROWPVT = Receptive One-Word Picture Vocabulary Test - 2, EOWPVT = Expressive One-Word Picture Vocabulary Test - 3, OWLS = Oral and Written Language Scales, RIST = Reynolds Intellectual Screening Test.

** Significant difference between groups, $p < 0.001$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Characteristics of the stimuli for analysis

		Rare phonotactic probability		Common phonotactic probability	
		Sparse n = 27	**Dense n = 23**	**Sparse n = 20**	**Dense n = 31**
Manipulated Variables					
Positional segment sum[a]	*M* (*SD*)	0.16 (0.06)	0.17 (0.03)	0.25 (0.10)	0.26 (0.06)
Biphone sum[a]	*M* (*SD*)	0.008 (0.004)	0.010 (0.005)	0.018 (0.012)	0.022 (0.009)
Neighborhood density[a]	*M* (*SD*)	5 (4)	8 (6)	4 (5)	9 (8)
Control Variables					
Age-of-acquisition[b]	*M* (*SD*)	3.9 (1.1)	3.8 (1.1)	3.7 (1.1)	3.8 (1.0)
Log word frequency[a]	*M* (*SD*)	2.6 (0.7)	2.6 (0.7)	2.7 (0.7)	2.6 (0.7)
Word length	*M* (*SD*)	4 (1)	4 (1)	4 (1)	4 (1)
Semantic set size [c]	*M* (*SD*)	13 (5)	12 (5)	13 (5)	14 (6)
Concreteness [d]	*M* (*SD*)	595 (36)	599 (26)	601 (14)	599 (21)
Familiarity [d]	*M* (*SD*)	533 (55)	511 (86)	541 (53)	533 (54)
Imagability [d]	*M* (*SD*)	595 (28)	596 (31)	599 (26)	596 (35)

[a]Computed from (Storkel et al., 2008).

[b]Computed from (Carroll & White, 1973; Garlock, 1997; Snodgrass & Yuditsky, 1996) where a rating of 1 corresponds to 0-2 years, rating of 5 corresponds to 6 years, and rating of 9 corresponds to 13+ years.

[c]Computed from (Nelson et al., 1998).

[d]Computed from (Wilson, 1987).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Correlations between demographic variables (columns) and experimental variables (rows)

Probe Type	Phonotactic Probability	Neighborhood Density	Chronological Age	ROWPVT raw score	EOWPVT raw score
Expressive	Rare	Sparse	0.44**	0.66**	0.59**
		Dense	0.65**	0.77**	0.68**
	Common	Sparse	0.60**	0.67**	0.60**
		Dense	0.45**	0.62**	0.57**
Receptive	Rare	Sparse	0.48**	0.53**	0.43**
		Dense	0.56**	0.59**	0.47**
	Common	Sparse	0.47**	0.52**	0.44**
		Dense	0.45**	0.54**	0.43**

Note. ROWPVT = Receptive One-Word Picture Vocabulary Test - 2, EOWPVT = Expressive One-Word Picture Vocabulary Test - 3.

** Significant correlation, $p < 0.01$

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 4

Means (and standard deviations) for children with phonological delays (PD) and children with typical development (TD) by phonotactic probability (rare vs. common) and neighborhood density (sparse vs. dense) for each probe type (expressive vs. receptive)

	Children with PD				Children with TD			
	Rare		Common		Rare		Common	
	Sparse	Dense	Sparse	Dense	Sparse	Dense	Sparse	Dense
Expressive	0.63[ac]	0.61[c]	0.58[a]	0.58	0.63[ac]	0.60[c]	0.58[a]	0.61
	(0.13)	(0.14)	(0.13)	(0.13)	(0.08)	(0.12)	(0.12)	(0.09)
Receptive	0.88[c]	0.86[c]	0.89	0.87	0.88[c]	0.83[bc]	0.88	0.89[b]
	(0.08)	(0.08)	(0.11)	(0.08)	(0.09)	(0.07)	(0.08)	(0.06)

[a]Effect of phonotactic probability in sparse neighborhoods: Rare significantly more accurate than common for both groups on the expressive probe.

[b]Effect of phonotactic probability in dense neighborhoods: Children with TD significantly more accurate for common than rare on the receptive probe.

[c]Effect of neighborhood density in rare sound sequences: Sparse significantly more accurate than dense for both groups on both probes.