RESEARCH ARTICLE

Word learning by children with phonological delays: Differentiating effects of phonotactic probability and neighborhood density

Holly L. Storkel

Jill R. Hoover

University of Kansas

In Press, *Journal of Communication Disorders*, 11/12/2009

Corresponding 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
E-mail: hstorkel@ku.edu
Phone: 785-864-0497
Fax: 785-864-3974

Abstract

This study examined the ability of 20 preschool children with functional phonological delays and 34 age- and vocabulary-matched typical children to learn words differing in phonotactic probability (i.e., the likelihood of occurrence of a sound sequence) and neighborhood density (i.e., the number of words that differ from a target by one phoneme). Children were exposed to nonwords paired with novel objects in a story and learning was measured by a picture naming task. Results showed that both groups created lexical representations for rare sound sequences from sparse neighborhoods. However, only children with typical development appeared to build on this initial lexical representation to create a full representation of the word (i.e., lexical-semantic connection and semantic representation). It was hypothesized that creating a lexical representation may be too resource demanding for children with phonological delays, leaving few resources available to create a lexical-semantic connection and/or a semantic representation.

Learning outcomes

The reader will be able to (1) define phonotactic probability; (2) define neighborhood density; (3) identify how these variables impact the word learning process in general; (4) identify potential areas of deficit in the word learning process for children with functional phonological delays.

1. Introduction

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). A full understanding of the nature of this disorder, in terms of the language representations and processes affected, has remained elusive. Hypothesized deficits include poor speech perception, poor oral-motor control as revealed by acoustic and kinematic measures, and poor higher level phonological knowledge, such as understanding how sounds are used to contrast meaning as well as how sounds can be combined to create words (see Munson, Edwards, & Beckman, 2005a for review). Thus, hypothesized deficits focus on deficits to motor and/or *phonological representations*, specifically representations of individual sounds in long-term memory. Moreover, the impact of these different hypothesized deficits on other areas of language acquisition has been relatively unexplored. It generally has been assumed that other areas of language are intact in children with functional phonological delays. However, more recent research has suggested that this assumption may be false. One language area that may be affected by phonological delay is word learning. Word learning involves *lexical representations*, the representation of the whole-word sound form in long-term memory, and *semantic representations*, the representation of the meaning or referent of a word in long-term memory.

To learn a word, a child must first recognize that a novel word was encountered, thereby triggering learning processes. It has been hypothesized that two characteristics contribute to this aspect of word learning. The first is the novelty of the word in the language as measured by *phonotactic probability*, the likelihood of occurrence of a sound sequence. That is, words that are more unique are more likely to be identified as novel, triggering learning processes (Storkel, Armbruster, & Hogan, 2006). Specifically, rarer sound sequences trigger learning more efficiently than more common sound sequences (Storkel et al., 2006). The second characteristic is similarity to other known words, termed *neighborhood density* for phonological similarity or *semantic set size* for semantic similarity. Here, presentation of a word activates representations of known words in long-term memory. For a novel word, none of the existing lexical or semantic representations in long-term memory will exactly match the novel word. This mismatch between the input and the child's representations in long-term memory is thought to trigger learning processes (Storkel & Adlof, 2009). When a novel word is similar to few other known words, as in a sparse neighborhood or small set size, the mismatch will be greater than when a novel word is similar to many other known words, as in a dense neighborhood or large set size, thereby facilitating initiation of learning (Storkel & Adlof, 2009). Once learning is initiated, the child must create a lexical and semantic representation of the word in long-term memory. This does not mark the end of word learning. Rather, the new lexical and semantic representations in long-term memory must form connections with existing lexical and semantic representations. This period of integration appears to occur separately from the creation of the representation and may be more protracted (Gaskell & Dumay, 2003; Leach & Samuel, 2007). Moreover, forming connections with many existing representations, as in a dense neighborhood or large set size, may serve to strengthen the new representation relative to forming connections with few existing representations, as in a sparse neighborhood or small set size (Storkel et al., 2006).

What is known about word learning by children with phonological delays? Edwards, Fox and Rogers (2002) provided evidence that children with phonological delays are less accurate discriminating words differing by a final consonant than children with typical phonological development (see also Edwards, Fourakis, Beckman, & Fox, 1999). Interestingly, this deficit was not tied to a specific error pattern in production (i.e., children showed this difficulty regardless of their production accuracy for final consonants). Poor discrimination could impact word learning by affecting the ability to identify that a novel word does not exactly match any existing lexical representations in long-term memory, resulting in a failure to trigger learning. Alternatively, poor discrimination may lead to misperception of the novel sound sequence, leading to inaccuracies in the newly created lexical representation. Importantly, Edwards and colleagues provide initial support for a relationship between speech perception and word learning. Specifically, they found a relationship between discrimination accuracy and vocabulary size (as well as articulatory accuracy). Moreover, previous research has demonstrated that children with phonological delays have lower receptive and expressive vocabulary scores on standardized tests than children with typical phonological development and that this difference persists into adulthood even after the production deficit has apparently resolved (Felsenfeld, Broen, & McGue, 1992; see also Shriberg & Kwiatkowski, 1994 for similar child findings).

Storkel (2004a) provided a more detailed picture of word learning by children with phonological delays. In this study, children with phonological delays and children with typical development learned nonwords that varied in phonotactic probability/neighborhood density. Phonotactic probability is positively correlated with neighborhood density in English (Storkel, 2004b; Vitevitch, Luce, Pisoni, & Auer, 1999). Specifically, rare sound sequences tend to have few neighbors (i.e., sparse neighborhoods) and common sound sequences tend to have many neighbors (i.e., dense neighborhoods). Thus, children in Storkel (2004a) were exposed to rare sound sequences from sparse neighborhoods (e.g., /gaʊb/) and common sound sequences from dense neighborhoods (e.g., /mæb/). Results of Storkel (2004a) showed differing effects of correlated phonotactic probability/neighborhood density based on phonological development.

In Storkel (2004a), children with phonological delays learned novel words composed of rare sound sequences in sparse neighborhoods more readily than novel words composed of common sound sequences in dense neighborhoods. In contrast, children with typical development showed the opposite pattern, learning novel words composed of common sound sequences in dense neighborhoods more readily than those composed of rare sound sequences in sparse neighborhoods (Storkel, 2004a). Moreover, error analyses provided evidence that common-dense sound sequences were particularly difficult to learn for the children with phonological delays. Specifically, Storkel (2004a) analyzed phonological errors, namely responses that shared 1 or more phonemes with the target nonword (e.g., [meɪp] – candy machine for target /mæb/ – pet, Storkel, 2004a, p. 1199), semantic errors, defined as responses that shared semantic category with the target nonword (e.g., [meɪp] – *candy machine* for target /ɡif/ – *candy machine,* Storkel, 2004a, p. 1199), and unrelated errors, specifically responses that did not share phonemes or semantic category with the target nonword (e.g., [meɪp] – candy machine for target /g□ɪt/ – pet, Storkel, 2004a, p. 1199). It was hypothesized that phonological and semantic errors indicated an emerging lexical or semantic representation respectively, characterized by partial knowledge of the target word (e.g., knowledge of the first sound or semantic category); whereas unrelated errors indicated impoverished lexical and/or semantic representations, characterized by limited, if any, knowledge of the target word (e.g., no knowledge of target sounds or semantic features). Results showed that children with phonological delays infrequently made phonological or semantic errors, but frequently made unrelated errors, when naming common-dense targets. In contrast, children with phonological delays produced all error types (i.e., phonological, semantic, unrelated) when naming rare-sparse targets. Moreover, children with phonological delays infrequently produced common-dense sound sequences as substitutes for other targets, producing rare-sparse sound sequences instead. From these patterns, it was hypothesized that children with phonological delays had impoverished representations of common sound sequences in dense neighborhoods, whereas they had a range of representations from impoverished to emerging to correct for rare sound sequences in sparse neighborhoods. In terms of patterns for the children with typical development, they produced all error types (i.e., phonologic, semantic, and unrelated) with equal frequency for both word types (i.e., rare-sparse, common-dense), indicating a range of representations.

One potential explanation of these results is that children with phonological delays had difficulty creating accurate and complete lexical representations (i.e., a lexical representation that exactly matches the target) for words in dense neighborhoods due to difficulty discriminating similar sound sequences (as also shown in Edwards et al., 1999; Edwards et al., 2002), whereas children with typical phonological development did not have this same problem and thus benefited from the multiple connections between new and existing lexical representations inherent in dense neighborhoods. One shortcoming of this hypothesis is that it implies that the results are attributable to neighborhood density, even though phonotactic probability also varied. A recent study of adult word learning showed independent effects of phonotactic probability and neighborhood density (Storkel et al., 2006) that may have implications for understanding word learning by children with phonological delays.

In Storkel et al. (2006), adults learned nonwords varying in both phonotactic probability and neighborhood density (i.e., rare-sparse, rare-dense, common-sparse, and common-dense). The effects of these two variables were examined in partially correct responses (i.e., 2 of 3 phonemes correct), which were thought to index early stages of word learning (i.e., initiating word learning), and fully correct responses (i.e., 3 of 3 phonemes), which were thought to index later stages of word learning (i.e., formation and integration of a new representation). Results showed that phonotactic probability influenced partially correct responses, whereas neighborhood density influenced completely correct responses. For phonotactic probability, adults produced more partially correct responses for rare sound sequences than for common sound sequences. It was thought that rare sound sequences triggered word learning more efficiently than common sound sequences. For neighborhood density, adults produced more completely correct responses for nonwords in dense neighborhoods than for nonwords in sparse neighborhoods. It was hypothesized that the greater number of connections between the newly created lexical representation and existing lexical representations in a dense neighborhood strengthened the newly created lexical representation, speeding word learning compared to the relatively fewer connections available in a sparse neighborhood. These findings suggest that differentiating the effect of phonotactic probability from that of neighborhood density in word learning by children with phonological delays may provide a clearer picture of word learning abilities in this population.

The purpose of the current study was to differentiate effects of phonotactic probability from those of neighborhood density in word learning by children with phonological delays and by children with typical development. The methods directly parallel those used in the previously described study of adult word learning (Storkel et al., 2006), including the examination of partially correct versus completely correct responses to investigate the influence of these two variables on early versus later stages of word learning. In addition, error patterns (i.e., semantic errors, unrelated errors) also were analyzed as in the earlier study of word learning by children with phonological delays (Storkel, 2004a) because this approach previously revealed that certain words, namely common sound sequences in dense neighborhoods, were particularly difficult to learn for children with phonological delays, who seemed to acquire only impoverished representations of these words.

2. Methods

2. 1 Participants

Fifty-four preschool children (age 3; 5 – 6; 7) participated: 20 with functional phonological delays and 34 with typical development. None of the children had a history of social, emotional, cognitive, motor, visual, hearing, or major medical impairments by parent report. All children passed a hearing screening in both ears (ASHA, 1997). Other characteristics of both groups of participants are shown in Table 1. Two main criteria were used to diagnose children with functional phonological delays. The first criterion related to defining a significant phonological delay. This was determined by administering the *Goldman-Fristoe Test of Articulation – 2* (GFTA, Goldman & Fristoe, 2000) and an extensive probe of English phonology (Gierut, 2008). Children either scored at the 11th percentile or below on the GFTA (n = 17) or scored between the 12th and 14th percentile on the GFTA and had at least six target sounds with inventory or positional constraints (n = 3) based on phonological analysis of both samples. The second criterion related to defining typical development in other areas. Here, children scored at or above the 16th percentile (1 standard deviation below the mean) on standardized tests of receptive vocabulary, expressive vocabulary, omnibus receptive language, and nonverbal intelligence (Brownell, 2000a, 2000b; Carrow-Woolfolk, 1995; Reynolds & Kamphaus, 2003).

Children with typical development scored at or above the 24th percentile on the GFTA and scored at or above the 16th percentile on standardized tests of receptive and/or expressive vocabulary (Brownell, 2000a, 2000b). In addition, this group was matched in gender, age, and raw receptive vocabulary scores to the group with phonological delays. T test comparisons showed that the groups differed significantly in their percentile ranks on the GFTA, t (52) = 9.05, $p < 0.001$, but not in chronological age or vocabulary test scores, all t (52) < 0.90, all $p > 0.35$. Gender also did not differ significantly between the groups, χ^2 (54) = 0.04, $p > 0.80$.

2.2 Stimuli

The same procedures for selecting nonwords used in Storkel et al. (2006) were used in this study. The exact same novel objects, stories, measures of learning, and procedures used in Storkel et al. (2006) were used in this study with one minor change: the number of exposures to the stimuli was increased in this study to guard against floor effects. Nonwords are described in detail here. Repeated stimuli and procedures are described briefly below with more detailed examples available in the previous publication.

2.2.1 Nonwords

Sixteen consonant-vowel-consonant (CVC) nonwords composed of early acquired consonants (i.e., glides, anterior nasals, and anterior stops) were selected to manipulate two independent variables: phonotactic probability and neighborhood density. Phonotactic probability and neighborhood density were computed following the same procedures as in Storkel et al. (2006) and using the same 20,000 word English corpus. Characteristics of the selected stimuli are shown in Table 2. Two measures of phonotactic probability were computed: positional segment sum and biphone sum. *Positional segment sum* is computed by adding the positional segment frequency of each sound in the nonword. Positional segment frequency is computed by adding the log frequency of every word in the dictionary containing the same sound in the same word position and then dividing by the sum of the log frequency of every word in the dictionary containing any sound in the same word position. *Biphone sum* is computed by adding the biphone frequency of each pair of sounds in the nonword. Biphone frequency is computed by adding the log frequency of every word in the dictionary containing the same pair of sounds in the same word position and then dividing by the sum of the log frequency of every word in the dictionary containing any sound in the same word position. These values were computed for all legal CVC nonwords containing early acquired consonants, and then a median split was used to classify the nonwords as having rare versus common sound sequences.

Neighborhood density was computed by counting all the words in the dictionary that differed from a given CVC nonword by a single sound substitution, addition, or deletion in any word position. As with phonotactic probability, neighborhood density was computed for all legal CVC nonwords containing early acquired consonants, and then a median split was used to classify the nonwords as residing in sparse versus dense neighborhoods.

After the stimuli were selected, an on-line child calculator became available (http://www.bncdnet.ku.edu/cgi-bin/DEEC/post_ccc.vi) that used these same algorithms to calculate positional segment sum, biphone sum, and neighborhood density using kindergarten and first grade child corpora (Kolson, 1960; Moe, Hopkins, & Rush, 1982). In general, the child values resulted in a similar classification of the stimuli (see Table 2).

2.2.2 Novel objects

The 16 novel objects used by Storkel et al. (2006) and Storkel (2004a) were used in this study. This set consisted of four novel objects in each of four semantic categories (i.e., toys, pets, candy machines, and horns). As in Storkel et al. (2006), one novel object from each semantic category was paired with a nonword from each phonotactic probability-neighborhood density condition, and pairings of novel objects and nonwords were counterbalanced across participants (refer to Table 3 in Storkel et al., 2006 for further details).

2.2.3 Stories

The two stories used by Storkel et al. (2006) were used in this study with two alterations: the script was adjusted to (1) increase the number of exposures to the nonwords to guard against floor effects and (2) include one direct imitation attempt for each nonword to judge whether children could produce the nonwords as intended. Each story consisted of three distinct story episodes. Each episode contained two novel objects from each of the four semantic categories, providing exposure to half of the novel object – nonword pairs. Each story episode provided four exposures to each novel object – nonword pair assigned to that story. This is an increase in exposure from Storkel et al. (2006) where the first episode provided only one exposure to the novel object – nonword pairs and the remaining two episodes provided three exposures to the novel object – nonword pairs (see the appendix of Storkel et al., 2006 for a sample story episode). Thus, the current stories provided 12 exposures to the novel object – nonword pairs after all three episodes were administered compared to 7 exposures in Storkel et al. (2006). The order of presentation of the two stories was counterbalanced across participants.

Additional exposures and a direct imitation attempt were provided via a review exposure. The review exposure was presented after each story episode. In the review exposure, each novel object from the story was presented in random order with the following script: “Look it’s a *nonword*. Say *nonword*.” Child produces nonword. “Remember, it’s a *nonword*.” In this way, the review exposure provided an additional four exposures to the novel object – nonword pairs, including a production attempt. All children imitated the stimuli as intended. In summary, after all three story episodes and reviews were presented, children had accrued 24 exposures to the novel object – nonword pairs.

Stories and review exposures were audio recorded by a female speaker in a sound proof booth, digitized, and edited. Two unfamiliar listeners trained in phonetic transcription listened to the recordings under the same presentation conditions as the participants to verify that all nonwords were presented as intended.

2.2.4 Measure of learning

Learning was measured using the same picture naming task described in Storkel et al. (2006). Each novel object was presented and the participant attempted to recall and produce its nonword label. Responses were phonetically transcribed and scored in five ways as a means of evaluating the status of newly created lexical and semantic representations in long-term memory. Inferring the status of lexical and semantic representations from production data is in line with methods used in past word learning studies (e.g., Gray, 2004; Storkel, 2004a; Storkel et al., 2006) and past naming studies (e.g., McGregor, 1997; McGregor & Appel, 2002; McGregor, Friedman, Reilly, & Newman, 2002; McGregor, Newman, Reilly, & Capone, 2002). A nonword was scored as completely correct if the participant produced all of the target phonemes in the correct order. Completely correct responses were thought to be indicative of a complete and accurate representation of the new word.

A nonword was scored as partially correct if the participant produced two of the three target phonemes in the correct order. Partially correct responses were thought to be indicative of partially complete representation of the new word (i.e., one that included specification of two of the three target phonemes as well as at least partial specification of the referent).

In the case where a response was neither partially correct nor completely correct, the error was coded further as semantic, unrelated, or unscorable. A semantic error was scored if the child's response contained two or three phonemes of another nonword presented in the study and if the non-target nonword had been paired with a novel object from the same semantic category as the target nonword (e.g., [haʊd] -- toy 2 for target /paɪb/ -- toy 1). A semantic error would likely occur if the child had an emerging semantic representation (e.g., semantic representation of the category but lacking semantic detail to differentiate it clearly from other members of the category).

An unrelated error was scored if the child's response contained two or three phonemes of another nonword presented in the study and if the non-target nonword had been paired with a novel object from a different semantic category than the target nonword (e.g., [woʊn] -- horn 1 for target /paɪb/ -- toy 1). An unrelated error might occur if the child has a relatively impoverished semantic and/or lexical representation (e.g., semantic representation containing no knowledge of target semantic features and/or lexical representation containing no knowledge of target sounds). In the case of unrelated errors, the phonotactic probability and density classification of the substituted nonword was noted (e.g., common-dense for /woʊn/) and analyzed. It was hypothesized that analysis of substitutes would reveal the status of the lexical representation of the substituted nonword. That is, the nonword that is produced as a substitute has, at least, a partially complete lexical representation to support production (e.g., a lexical representation that specified at least two of the target phonemes), although other representations (e.g., semantic) could be impoverished.

A response was considered unscorable if it did not match any of the nonwords in the study. Thus, unscorable responses included invented nonwords, real words, and no responses.

Unscorable responses were not analyzed further. Note that real words could be semantically related to the nonobject (e.g., “toy” for target /paɪb/ -- toy 1). However, these were coded as unscorable rather than a semantic error because many semantic features could be inferred simply by looking at the nonobject. Thus, these errors do not necessarily indicate that the child has learned something about the nonobject. A person who had not undergone the training might produce the same response. In fact, this is the type of response children provided during baseline testing prior to training.

Consonant-to-consonant transcription reliability was computed for transcription of the real words produced by the child on the *Goldman-Fristoe Test of Articulation-2*, English phonology probe, and the nonwords produced during the word learning task (i.e., repetition and naming) for 19% of the participants. Reliability was similar for each group so aggregate reliability is reported here. Interjudge transcription reliability was 97% (*SD* = 2%, range 94-99%) for real words and 98% (*SD* = 2%, range 94-100%) for nonwords.

Scoring reliability was computed for the picture naming test for 24% of the participants. Reliability was similar for each group so aggregate reliability is reported here. Two types of scoring reliability were computed. The first examined classification of the child’s response as a particular nonword. Interjudge reliability for this score was 98% (*SD* = 2%, range = 91% - 100%). The second examined classification of the child’s response relative to the target word as partially correct, fully correct, semantic error, unrelated error, or unscorable. Interjudge reliability for this score was 98% (*SD* = 4%, range = 84% - 100%).

2.3 Procedures

The word learning protocol required three sessions. For each session, the participant was seated in front of a computer connected to table top speakers. Delivery of the audio and visual stimuli was controlled by Direct RT software (Jarvis, 2002). Children wore a head-mounted microphone connected via a splitter to audio and video recording equipment.

The first session focused on the first story. The session began with baseline testing. Each novel object from the first story was presented individually on the computer screen and children were encouraged to guess its name. The order of presentation of the novel objects was randomized by the experimental control software (Jarvis, 2002). Then, the first story episode was administered followed by the review exposure. The picture naming task was then re-administered but with altered instructions to encourage the children to recall the names for the novel objects from the story. The second and third story episodes, corresponding reviews, and picture naming tests were administered following the same procedures. The entire session lasted approximately 40 minutes.

The second session, which occurred approximately 1-week later (*M* = 7 days, *SD* = 2 days, range 2-14 days), began with the review of the stimuli from the first story followed by a picture naming test for the first story. The remainder of the session focused on the stimuli from the second story, which were administered following the same procedures described for the first session. The entire session lasted approximately 45 minutes.

The third session, which occurred approximately 1-week later (*M* = 7 days, *SD* = 2 days, range 2 – 14 days), began with the review of the stimuli from the second story followed by a picture naming test for the second story. The remainder of the session consisted of administration of clinical tests. Additional sessions occurred either prior to or after administration of the word learning protocol to complete the clinical test battery described in the participants section.

Procedural reliability for administration of the word learning protocol was computed for 20% of the participants. Reliability was similar for each group so aggregate reliability is reported here. A reliability judge viewed the video tape for the selected participants and scored protocol administration (e.g., correct version of protocol administered, tasks administered in the correct order, correct directions and feedback provided, correct on-line scoring) as well as equipment and set-up (e.g., appropriate audio and video quality, appropriate speaker loudness, computer malfunctions, data collection forms completed correctly). Procedural reliability was 96% (*SD* = 4%, range 87-100%).

3. Results

There were four dependent variables related to accuracy and errors: proportion of completely correct responses (i.e., 3 of 3 phonemes correct), proportion of partially correct responses (i.e., 2 of 3 phonemes correct), proportion of semantic error responses, and proportion of unrelated error responses. There was an additional dependent variable specific to unrelated errors, proportion of unrelated errors, which was used to examine the phonotactic probability and neighborhood density of the nonwords that were produced as substitutes for unrelated targets. Each dependent variable was analyzed separately. Each analysis used a 2 Group (children with phonological delay vs. children with typical development) x 2 Phonotactic Probability (rare vs. common) x 2 Neighborhood Density (sparse vs. dense) x 2 Time (24-exposure vs. 1-week post-exposure) mixed ANOVA. Note that not all time periods were analyzed (e.g., cycle 1/8-exposures, and cycle 2/16-exposures) due to overall low performance at these earlier test points.

3.1 Completely correct responses

For completely correct responses, no main effects or interactions that addressed the research questions were statistically significant, all *F*s < 4.00, all *ps* > 0.05, all $\eta_p^2 < 0.08$.

3.2 Partially correct responses

For partially correct responses, there was a significant two-way interaction of group and density, F (1, 52) = 4.22, $p < 0.05$, $\eta_p^2 = 0.08$., which was qualified by a significant three-way interaction of group, neighborhood density, and phonotactic probability, F (1, 52) = 4.02, p = 0.05, $\eta_p^2 = 0.07$. These interactions were explored by examining the effect of group and neighborhood density at each level of phonotactic probability (rare vs. common), using a 2 Group (children with phonological delay vs. children with typical development) x 2 Neighborhood Density (sparse vs. dense) x 2 Time (24-exposure vs. 1-week post-exposure) mixed ANOVA.

3.2.1 Rare sound sequences

For rare sound sequences, there was a significant interaction between group and density, F (1, 52) = 9.11, $p < 0.01$, $\eta_p^2 = 0.15$. As shown in Figure 1, children with phonological delays tended to produce more partially correct responses to nonwords from dense neighborhoods (M = 0.06, SE = 0.02, SD = 0.11) than to nonwords from sparse neighborhoods (M = 0.03, SE = 0.02, SD = 0.10), although this effect failed to achieve statistical significance despite a medium effect size, F (1, 19) = 3.35, p = 0.08, $\eta_p^2 = 0.15$. In contrast, children with typical development produced significantly more partially correct responses to nonwords from sparse neighborhoods (M = 0.08, SE = 0.02, SD = 0.15) than to nonwords from dense neighborhoods (M = 0.02, SE = 0.01, SD = 0.08), F (1, 33) = 7.65, $p < 0.01$, $\eta_p^2 = 0.19$.

3.2.2 Common sound sequences

For common sound sequences, no main effects or interactions that addressed the research questions were statistically significant, all Fs < 1.65, all ps > 0.20, all $\eta_p^2 < 0.04$. As shown in Figure 1, neighborhood density did not appear to influence learning of common sound sequences for either children with phonological delays or children with typical development.

3.2.3 Additional group comparison

Analysis of rare sound sequences suggested potential group differences in the effect of neighborhood density. However, the finding from the children with phonological delays only approached significance. Group differences were further confirmed by directly comparing children with phonological delays to those with typical development. This analysis showed group differences for nonwords from sparse neighborhoods. Specifically, children with typical development ($M = 0.07$, $SE = 0.01$, $SD = 0.06$) produced more partially correct responses than children with phonological delays ($M = 0.03$, $SE = 0.02$, $SD = 0.07$) for nonwords in sparse neighborhoods, $F\ (1, 52) = 4.62$, $p = 0.04$, $\eta_p^2 = 0.08$.

3.3 Semantic errors

For semantic errors, there were no significant main effects or interactions, Fs < 3.10, all $ps > 0.08$, all $\eta_p^2 < 0.06$.

3.4 Unrelated errors

For unrelated errors, there was a significant three-way interaction of group, phonotactic probability, and neighborhood density, $F\ (1, 52) = 6.09$, $p < 0.05$, $\eta_p^2 = 0.11$. No other main effects or interactions were significant, all Fs < 3.50, all $ps > 0.05$, all $\eta_p^2 < 0.07$. The significant three-way interaction was further explored by examining the effect of group and neighborhood density at each level of phonotactic probability (rare vs. common), using a 2 Group (children with phonological delay vs. children with typical development) x 2 Neighborhood Density (sparse vs. dense) x 2 Time (24-exposure vs. 1-week post-exposure) mixed ANOVA.

3.4.1 Rare sound sequences

For rare sound sequences, there were no significant main effects or interactions, *F*s < 1.90, all *ps* > 0.17, all $\eta_p^2 < 0.04$. Neighborhood density and phonological development did not appear to affect production of unrelated errors (see Figure 2).

3.4.2 Common sound sequences

For common sound sequences, there was a significant interaction between group and density, F (1, 52) = 5.94, $p < 0.05$, $\eta_p^2 = 0.10$. As shown in Figure 2, children with phonological delays produced significantly more unrelated errors to nonwords from sparse neighborhoods (*M* = 0.23, *SE* = 0.05, *SD* = 0.24) than to nonwords from dense neighborhoods (*M* = 0.14, *SE* = 0.03, *SD* = 0.14, F (1, 19) = 6.33, $p < 0.05$, $\eta_p^2 = 0.25$. In contrast, children with typical phonological development showed no effect of neighborhood density on unrelated errors, *F*s < 1.80, all *ps* > 0.18, all $\eta_p^2 < 0.06$.

3.5 Unrelated Substitutes

The phonotactic probability and neighborhood density of the substitute nonwords for unrelated responses was analyzed. Results showed a significant two-way interaction of phonotactic probability and time, F (1, 39) = 8.22, $p < 0.01$, $\eta_p^2 = 0.17$, which was qualified by a significant three-way interaction of group, phonotactic probability, and time, F (1, 39) = 4.37, $p < 0.05$, $\eta_p^2 = 0.10$. No other main effects or interactions were significant, all *F*s < 3.60, all *ps* > 0.05, all $\eta_p^2 < 0.09$. The significant two-way and three-way interactions were further explored by examining the effect of group and phonotactic probability at each level of time (24-exposure vs. 1-week post-exposure) using a 2 Group (children with phonological delay vs. children with typical development) x 2 Phonotactic Probability (rare vs. common) x 2 Neighborhood Density (sparse vs. dense) mixed ANOVA.

3.5.1 24-exposure

As shown in Figure 3, at the 24-exposure test, children produced rare sound sequences ($M = 0.31$, $SD = 0.37$, $SE = 0.05$) as unrelated substitutes significantly more frequently than common sound sequences ($M = 0.19$, $SD = 0.25$, $SE = 0.03$), $F(1, 42) = 6.80$, $p = 0.02$, $\eta_p^2 = 0.14$. In addition, children produced nonwords from sparse neighborhoods ($M = 0.31$, $SD = 0.35$, $SE = 0.05$) as unrelated substitutes significantly more frequently than nonwords from dense neighborhoods ($M = 0.19$, $SD = 0.28$, $SE = 0.04$), $F(1, 42) = 5.80$, $p = 0.02$, $\eta_p^2 = 0.12$. As shown in Figure 3, these two main effects were additive such that the most frequently produced unrelated substitutes were rare sound sequences from sparse neighborhoods.

3.5.2 1-week post-exposure

At the 1-week post-exposure test, there were no significant main effects or interactions, Fs < 1.80, all $ps > 0.17$, all $\eta_p^2 < 0.04$. Phonotactic probability, neighborhood density, and phonological development did not appear to affect production of unrelated substitutes at the post-test.

4. Discussion

The purpose of this study was to examine the individual and combined influence of phonotactic probability and neighborhood density on word learning by children varying in phonological development (i.e., delayed vs. typical). Only partially correct responses, unrelated errors, and unrelated substitutes showed significant effects that address the research questions, whereas completely correct responses and semantic errors failed to provide insights into the role of these variables in word learning. Across partially correct responses and unrelated errors, the overall finding was that phonotactic probability and neighborhood density interacted with each other. This interaction between phonotactic probability and neighborhood density differs from past word learning studies of adults (Storkel et al., 2006) which yielded only main effects of each variable and no interaction. This suggests that children require a convergence of phonotactic probability and neighborhood density for optimal word learning. Importantly, the “optimal convergence” appeared to differ across children with phonological delays and children with typical development, as evidenced by the obtained interactions involving group.

Let’s begin by examining the one point of similarity across the groups: the effect of phonotactic probability and neighborhood density on unrelated substitutes. Specifically, both groups of children produced rare sound sequences in sparse neighborhoods more frequently as unrelated substitutes than any other combination of phonotactic probability and neighborhood density. Recall that it was hypothesized that nonwords produced as unrelated substitutes for other target words had, at least, partially complete lexical representations, although other representations (e.g., semantic) could be impoverished. Thus, both groups appeared to have at least partially complete lexical representations for rare sound sequences from sparse neighborhoods as shown by their frequent production as unrelated substitutes. This pattern fits well with previous hypotheses concerning the role of phonotactic probability and neighborhood density in triggering learning. Specifically, rare sound sequences are hypothesized to stand out as being novel more than common sound sequences, leading to more efficient triggering of learning (Storkel et al., 2006). In addition, novel words that reside in sparse neighborhoods will activate fewer existing lexical representations than novel words in dense neighborhoods. Thus, the mismatch between the input and the child’s existing lexical representations will be more apparent for sparse than for dense neighborhoods, facilitating triggering of learning (Storkel & Adlof, 2009). Taken together, rare phonotactic probability and sparse neighborhoods provide converging cues that a new word is present, efficiently triggering creation of a new lexical representation. Importantly, this process appears to be similarly intact for children with phonological delays and for children with typical development.

Although children with phonological delays and those with typical development trigger learning of a lexical representation of rare sound sequences from sparse neighborhoods similarly, there appear to be differences in other aspects of word learning. In particular, children with typical development appear to be able to build on this initial lexical representation for rare sound sequences in sparse neighborhoods to create partially complete lexical-semantic connections and semantic representations, yielding more partially correct responses for rare-sparse sound sequences than rare-dense sound sequences. In this way, the effect of phonotactic probability and neighborhood density on partially correct responses by children with typical development is viewed as a by-product of efficiently triggering creation of a new lexical representation. That is, children with typical development immediately recognize that a rare sound sequence from a sparse neighborhood is a new word and immediately initiate creation of a new lexical representation. This rapid start in creating the new lexical representation presumably allows later exposures to the new word to be devoted to creation of other representations, such as the lexical-semantic connection and the semantic representation. Thus, for children with typical development, rare sound sequences and sparse neighborhoods converge to create optimal word learning conditions, presumably due to their hypothesized role in triggering word learning.

Children with phonological delays do not appear to achieve the same success as children with typical development in building on the initial lexical representation of a rare sound sequences from a sparse neighborhood. In particular, children with phonological delays produced significantly fewer partially correct responses than typically developing children for nonwords from sparse neighborhoods. In addition, children with phonological delays showed a trend for more partially correct responses to rare sound sequences from dense neighborhoods than rare sound sequences from sparse neighborhoods. Thus, the initial advantage in creating a, at least, partially complete lexical representation for rare sound sequences from sparse neighborhoods may actually disappear for children with phonological delays at later point in word learning. That is, children with phonological delays, like children with typical development, immediately recognize that a rare sound sequence from a sparse neighborhood is a new word and immediately trigger creation of a new lexical representation. However, this initial rapid start in creating the new lexical representation may come at the expense of other representations (i.e., lexical-semantic connection, semantic representation) that are needed to support a partially correct response, leading to the observed group differences.

There is another piece of evidence suggesting that children with phonological delays may experience difficulty with sparse neighborhoods. In particular, children with phonological delays produced more unrelated errors to common sound sequences from sparse neighborhoods than common sound sequences from dense neighborhoods. Recall that unrelated errors were thought to index impoverished lexical and/or semantic representations. Thus, children with phonological delays appeared to have more impoverished lexical and/or semantic representations of common sound sequences from sparse neighborhoods than common sound sequences from dense neighborhoods. We assume that part of the reason that common sound sequences from sparse neighborhoods have impoverished lexical and/or semantic representations for children with phonological delays relates to the previous explanation of triggering word learning. That is, common sound sequences from sparse neighborhoods would not trigger word learning as efficiently as rare sound sequences from sparse neighborhoods because there would be conflicting cues to the novelty of the nonword and thus the need for new learning. As a result, creation of a new lexical representation may not be triggered immediately, which would have subsequent consequences (as previously described). Moreover, common sound sequences from sparse neighborhoods would later encounter whatever process interferes with the formation of a lexical-semantic connection and/or semantic representation in sparse neighborhoods, ultimately yielding an impoverished lexical and/or semantic representation.

What possible mechanism could account for difficulties creating a lexical-semantic connection and/or a semantic representation for novel words in sparse neighborhoods by children with phonological delays? The past literature suggests two cognitive processes where sparse neighborhoods lead to slower or more inaccurate processing than dense neighborhoods in adults. The first is production where adults tend to produce real words from sparse neighborhoods more slowly than real words from dense neighborhoods (Vitevitch, 2002). Thus, it’s possible that children with phonological delays experience greater difficulty with novel words in sparse neighborhoods during the production measure of word learning but not during word learning itself. This seems somewhat unlikely because children with phonological delays do produce the novel words from sparse neighborhoods, just not in response to the correct target object, suggesting a problem with word learning rather than production.

The second past finding of poorer performance for sparse neighborhoods than dense neighborhoods comes from the working memory literature where nonwords from sparse neighborhoods are recalled more poorly than nonwords from dense neighborhoods (Roodenrys & Hinton, 2002; Thorn & Frankish, 2005). In this way, children with phonological delays may immediately trigger learning of novel words from sparse neighborhoods but then experience difficulty holding these sound sequences in working memory during exposure. They do succeed in holding the sound sequence in memory so that an accurate and complete lexical representation can be created. However, this process may be so resource demanding that there are no resources remaining to create an accurate and complete lexical-semantic connection and/or semantic representation. This same process could be at work for children with typical development in that holding a novel word from a sparse neighborhood in working memory would be more resource demanding than holding a novel word from a dense neighborhood in working memory. Presumably, the difference between the groups is one of degree. That is, the children with phonological delays have much greater difficulty holding the sound sequences from sparse neighborhoods in working memory than the children with typical development. There have been few studies of working memory in children with phonological delays and none of these have manipulated neighborhood density. In general, children with phonological delays appear to perform more poorly than typically developing children on working memory tasks (Munson, Edwards, & Beckman, 2005b), implicating capacity or efficiency limitations in working memory, but it is unclear whether this difference would be particularly exacerbated for words from sparse neighborhoods.

These findings are somewhat, although not completely, at odds with past findings (Storkel, 2004a). Recall that group differences were noted between children with phonological delays and those with typical development in the past study but the optimal learning condition for each group in the past study differs from that identified in the current study. In addition, the hypothesized area of deficit in children with phonological delays in the past study differs from that in the current study. Taking up the first issue, the difference in the direction of the effect of phonotactic probability and neighborhood density may be attributable to the independent manipulation of phonotactic probability and neighborhood density, which allowed each cue to be used in different combinations for different aspects of word learning. The findings in the current study from typically developing children replicate those in another study independently manipulating phonotactic probability and neighborhood density (Hoover, Storkel, & Hogan, 2009). That study hypothesized that typically developing children required a convergence of cues to learn new words, and that the cues that were convergent depended on the particular word learning process. For example, rare sound sequences and sparse neighborhoods provide convergent cues that a sound sequence is novel and needs to be learned, a hypothesis that also fits the results of the current study for both groups of children.

Turning to the second issue, the difference in the area of deficit for children with phonological delays across studies could be due to a change in training methods across studies. In the previous study, children received only auditory exposures to the novel words. In the current study, children also produced the novel words in direct imitation. A recent study of adult word learning suggests that adding production attempts to the training component enhances learning of a lexical representation but hinders the creation of connections between representations (Leach & Samuel, 2007). In this vein, we hypothesize that adding the production component altered the word learning process slightly leading to a change in the area of deficit for children with phonological delays. That is, producing the novel words in direct imitation may have facilitated creation of a lexical representation for children with phonological delays. However, success at creating a lexical representation came at a cost to creation of lexical-semantic connection and/or a semantic representation. This hypothesis potentially has interesting implications for understanding how to best introduce new words to children with phonological delays, warranting explicit testing.

5. Conclusion

Rare sound sequences and sparse neighborhoods appear to provide converging cues to the novelty of a sound sequence, immediately triggering word learning. Immediate triggering of word learning may lead to creation of a relatively accurate and complete lexical representation for children with phonological delays and children with typical development. Children with typical development appear to be able to capitalize on this well formed lexical representation to create at least partially complete and accurate lexical-semantic connections and semantic representations for rare sound sequences from sparse neighborhoods. In contrast, children with phonological delays demonstrate difficulty in this area. Because of working memory constraints, the process of creating a relatively accurate and complete lexical representation may be too resource demanding for children with phonological delays, leaving few resources available to create a lexical-semantic connection and/or a semantic representation.

Acknowledgments

Holly L. Storkel, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas. Jill R. Hoover, Child Language Doctoral Program, University of Kansas.

This research was supported by DC 006545, DC 008095, DC 009135, DC 00052, DC 05803, and HD 02528. The following individuals contributed to stimulus creation, data collection, data processing, and reliability calculations: Teresa Brown, Jenny Fox, Andrea Giles, Stephanie Gonzales, Nicole Hayes, Jennica Kilwein, Su-Yeon Lee, Junko Maekawa, Shannon Rogers, Josie Row, Mariza Rosales, Katie Shatzer, Maki Sueto, Allison Wade, and Courtney Winn.

Appendix A. Continuing education

1. What is phonotactic probability?

    a. The number of times a word occurs in a language
    b. The likelihood of occurrence of a sound sequence
    c. The age when a word was learned
    d. The number of words that differ from a word by one phoneme

2. What is neighborhood density?

    a. The number of times a word occurs in a language
    b. The likelihood of occurrence of a sound sequence
    c. The age when a word was learned
    d. The number of words that differ from a word by one phoneme

3. What characteristics of a word would aid a child in recognizing that a novel word was encountered, efficiently triggering learning of the word?

    a. Rare sound sequence and sparse neighborhood
    b. Rare sound sequence and dense neighborhood
    c. Common sound sequence and sparse neighborhood
    d. Common sound sequence and dense neighborhood

4. What characteristic of a word would strengthen a newly created representation?

    a. Rare sound sequence
    b. Common sound sequence
    c. Sparse neighborhood
    d. Dense neighborhood

5. Children with phonological delays may have difficulty creating a lexical-semantic connection and/or a semantic representation. True or False

Answer Key

1. B
2. D
3. A
4. D
5. True

References

ASHA. (1997). Guidelines for screening for hearing impairment-preschool children, 3-5 years. *Asha, 4*, IV-74cc - IV-74ee.

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.

Carrow-Woolfolk, E. (1995). *Oral and written language scales*. Circle Pines, MN: American Guidance Service, Inc.

Edwards, J., Fourakis, M., Beckman, M. E., & Fox, R. A. (1999). Characterizing knowledge deficits in phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*, 169-186.

Edwards, J., Fox, R., & Rogers, C. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy. *Journal of Speech and Hearing Research, 45*, 231-242.

Felsenfeld, S., Broen, P. A., & McGue, M. (1992). A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. *Journal of Speech and Hearing Research, 35*, 1114-1125.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89* (2), 105-132.

Gierut, J. A. (2008). Phonological disorders and the developmental phonology archive. In D. A. Dinnsen & J. A. Gierut (Eds.), *Optimality theory, phonological acquisition and disorders* (pp. 37-92). London: Equinox.

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

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47* (5), 1117-1132.

Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2009). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Manuscript submitted for publication*.

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.

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

McGregor, K. K. (1997). The nature of word-finding errors of preschoolers with and without word-finding deficits. *Journal of Speech, Language, and Hearing Research, 40*, 1232-1244.

McGregor, K. K., & Appel, A. (2002). On the relation between mental representation and naming in a child with specific language impairment. *Clinical Linguistics & Phonetics, 16* (1), 1-20.

McGregor, K. K., Friedman, R. M., Reilly, R., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45* (2), 332-346.

McGregor, K. K., Newman, R. M., Reilly, R., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45* (5), 998-1014.

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

Munson, B., Edwards, J., & Beckman, M. E. (2005a). Phonological knowledge in typical and atypical speech-sound development. *Topics in Language Disorders, 25* (3), 190-206.

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

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds intellectual screening test*. Lutz, FL: Psychological Assessment Resources, Inc.

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

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders. I: A clinical profile. *Journal of Speech and Hearing Research, 37* (5), 1100-1126.

Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B. (1986). Characteristics of children with phonological disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51*, 140-161.

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* (5), 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* (6), 1454-1468.

Storkel, H. L., & Adlof, S. M. (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language, and Hearing Research, 52*, 306-320.

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* (6), 1175-1192.

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* (4), 729-735.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735-747.

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* (1-2), 306-311.

Figure Captions

Figure 1. Mean proportion of partially correct responses by children with phonological delays (PD) and children with typical development (TD) for rare sound sequences from sparse neighborhoods (open bar), rare sound sequences from dense neighborhoods (horizontal stripes), common sound sequences from sparse neighborhoods (solid bar), and common sound sequences from dense neighborhoods (vertical stripes). Error bars indicate standard errors.

Figure 2. Mean proportion of unrelated responses by children with phonological delays (PD) and children with typical development (TD) for rare sound sequences from sparse neighborhoods (open bar), rare sound sequences from dense neighborhoods (horizontal stripes), common sound sequences from sparse neighborhoods (solid bar), and common sound sequences from dense neighborhoods (vertical stripes). Error bars indicate standard errors.

Figure 3. Mean proportion of nonwords produced as unrelated substitutes at 24 exposures. The phonotactic probability (x-axis) and neighborhood density (shading of bar) of the substituted nonword are indicated.

Table 1

Participant characteristics.

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

	85 – 116	
OWLS expressive standard score	97	N/A
	12	
	80 - 119	
RIST standard score	115	N/A
	20	
	89 - 155	

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, $p < 0.001$.

Table 2

Phonotactic probability and neighborhood density of the stimuli.

		Rare phonotactic probability		Common phonotactic probability	
		Sparse[1]	Dense[2]	Sparse[3]	Dense[4]
Characteristics Based on Adult Corpus					
Positional segment frequency	*M*	0.09	0.09	0.15	0.17
	(SD)	(0.01)	(0.01)	(0.03)	(0.02)
Biphone frequency	*M*	0.0017	0.0019	0.0073	0.0062
	(SD)	(0.0010)	(0.0004)	(0.0058)	(0.0029)
Neighborhood density	*M*	7	13	6	16
	(SD)	(1)	(1)	(1)	(1)
Characteristics Based on Child Corpus					
Positional segment frequency	*M*	0.11	0.12	0.16	0.18
	(SD)	(0.02)	(0.02)	(0.05)	(0.02)
Biphone frequency	*M*	0.0029	0.0029	0.0059	0.0074
	(SD)	(0.0018)	(0.0005)	(0.0033)	(0.0015)
Neighborhood density	*M*	6	9	6	13
	(SD)	(2)	(3)	(3)	(2)

Nonwords by condition: [1]/hub t□ɪm nib w□p/ [2]/wud jeɪm n□d haʊd/ [3]/mɛb paɪb hɑn jaʊn/ [4]/woʊn jæt nɪd paʊn/

Fig 1

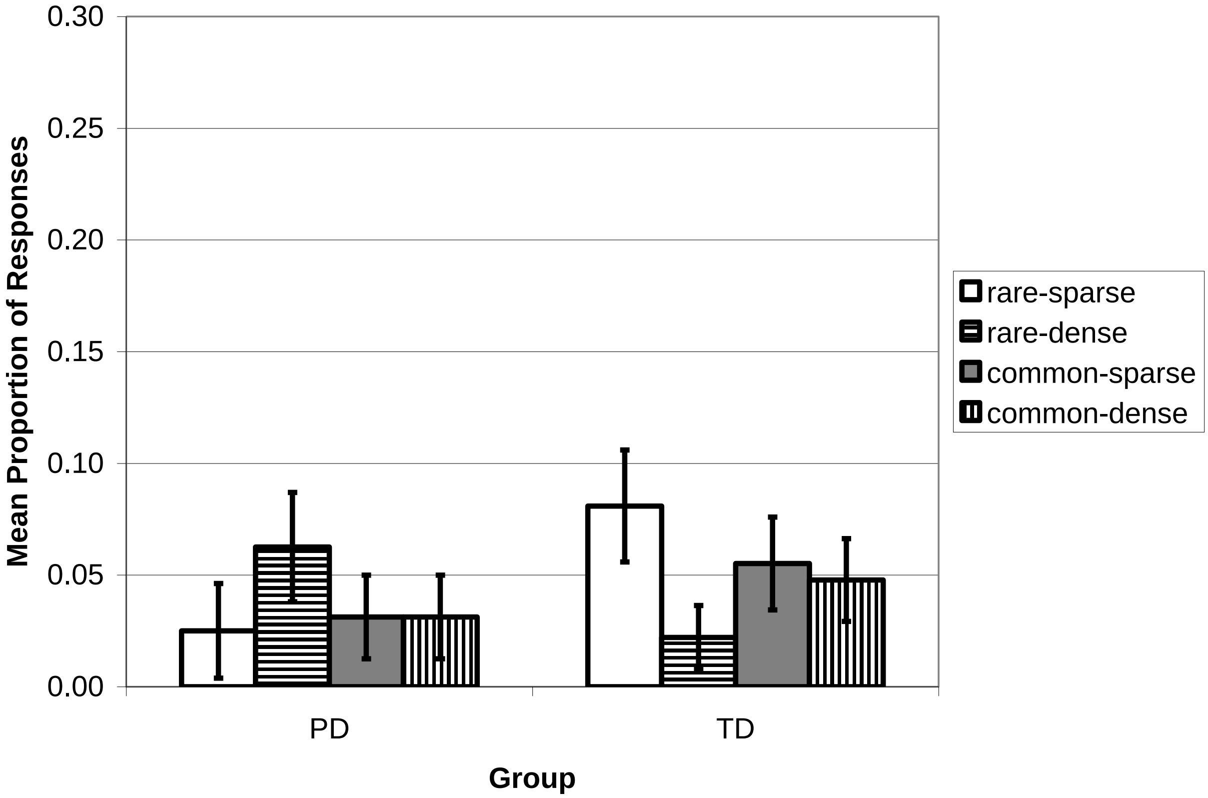

0.30
0.25
0.20
0.15
0.10
0.05
0.00
Mean Proportion of Responses
PD
TD
Group
rare-sparse
rare-dense
common-sparse
common-dense

Fig 2

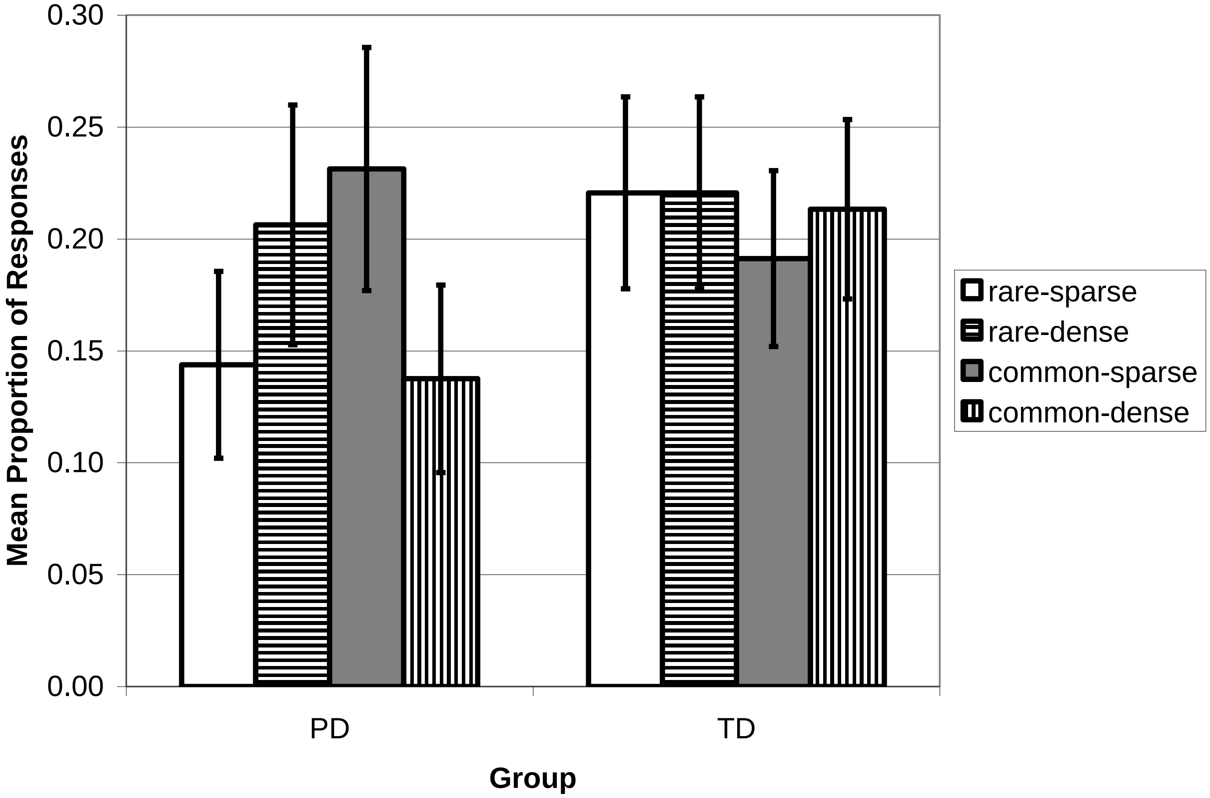
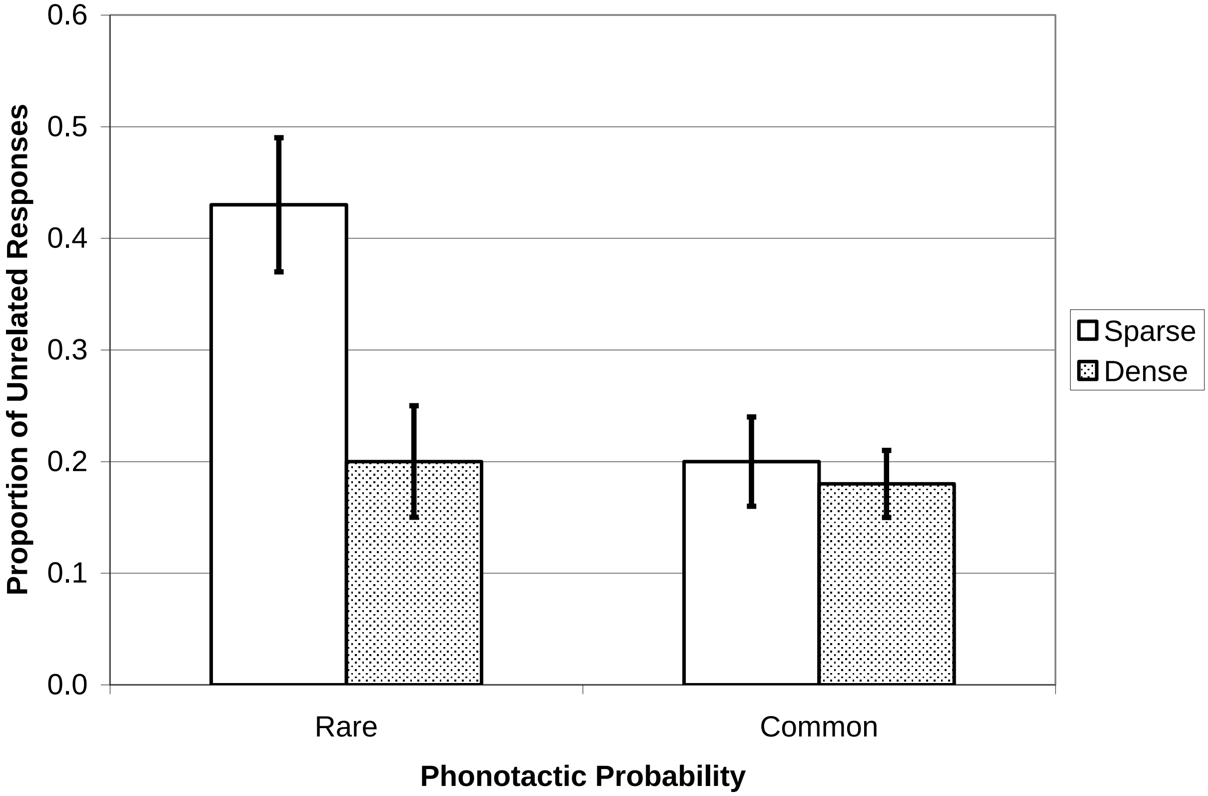

0.30
0.25
0.20
0.15
0.10
0.05
0.00
Mean Proportion of Responses
PD
TD
Group
rare-sparse
rare-dense
common-sparse
common-dense

Fig 3


0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion of Unrelated Responses
Rare
Common
Phonotactic Probability
☐ Sparse
▨ Dense