Published in final edited form as:
J Commun Disord. 2011 ; 44(1): 49–58. doi:10.1016/j.jcomdis.2010.07.002.

The influence of neighborhood density and word frequency on phoneme awareness in 2nd and 4th grades

Tiffany P. Hogan[1], **Ryan P. Bowles**[2], **Hugh W. Catts**[3], and **Holly L. Storkel**[3]
[1]University of Nebraska – Lincoln
[2]Michigan State University
[3]University of Kansas

Abstract

Purpose—The purpose of this study was to test the hypothesis that two lexical characteristics – neighborhood density and word frequency – interact to influence performance on phoneme awareness tasks.

Methods—Phoneme awareness was examined in a large, longitudinal dataset of 2nd and 4th grade children. Using linear logistic test model, the relation between words' neighborhood density, word frequency, and phoneme awareness performance was examined across grades while covarying type and place of deletion.

Results—A predicted interaction was revealed: words from dense neighborhoods or those with high frequency were more likely to yield correct phoneme awareness responses across grades.

Conclusions—Findings support an expansion to the lexical restructuring model to include interactions between neighborhood density and word frequency to account for phoneme awareness.

1. Introduction

Phoneme awareness assessment has become a critical part of early identification and diagnosis of reading disability. *Phoneme awareness*, one's sensitivity to the sound structure of a word, has an important role in early identification of reading impairment because of its causal link to early reading achievement (for reviews, see Adams, 1990; Gillon, 2004). Simply stated, children with poor phoneme awareness are more likely to become poor readers, whereas those with good phoneme awareness often become good readers. Although the relationship between phoneme awareness and reading has been explored extensively, less is known about the factors affecting phoneme awareness test performance.

Studies focusing on phoneme awareness test performance have examined issues such as task difficulty (Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993), picture support (Gibbs, 2003), and construct validity (Anthony & Lonigan, 2004). One area that has

Corresponding Author: Tiffany P. Hogan 359 Barkley Memorial Center University of Nebraska – Lincoln Lincoln, NE 68583 Phone: (402) 472-3911 Fax: (402) 472-7697 thogan2@unl.edu.

Portions of the research reported here were presented at the 2004 American Speech-Language-Hearing Association Annual Convention in Philadelphia, PA.

received less attention is the influence of *lexical characteristics*, the sound properties of words, on phoneme awareness performance. Lexical characteristics are predicted to influence phoneme awareness performance according to the lexical restructuring model (Metsala & Walley, 1998; Walley, Metsala, & Garlock, 2003). This study examines a longitudinal dataset of item level phoneme awareness performance to investigate tenets of the lexical restructuring model. Understanding factors associated with phoneme awareness performance will lead to more precise assessments and treatments for children at risk for reading disabilities.

1.1 The Lexical Restructuring Model

The lexical restructuring model hypothesizes that phoneme awareness is the product of the segmental restructuring of lexical representations that arises as a result of young children's rapid vocabulary growth (Fowler, 1991; Metsala & Walley, 1998; Walley et al., 2003). That is, when children's lexicons contain few words, holistic representations of those words, including minimal phonemic information, are sufficient to differentiate each word from every other word. As children expand their vocabularies, underlying lexical representations must become more phonemically detailed to differentiate newly learned words from existing words in the lexicon[1]. The lexical restructuring model states that this process of increased phonemic detail can extend into later childhood as children's lexicons expand. Words contain phonemic detail based on lexical characteristics that describe their place in the lexicon. Two lexical characteristics proposed to be associated with phonemic detail are neighborhood density and word frequency.

1.1.1 Neighborhood Density—Neighborhood density refers to the number of phonologically similar words in the lexicon, and is most often calculated by determining the number of words that are created by adding, deleting, or substituting a single sound in a given word (Luce & Pisoni, 1998). Thus, the word "sit" has 36 neighbors including "spit," "it," and "hit". Words with a high number of neighbors are said to reside in dense neighborhoods, whereas those with few neighbors reside in sparse neighborhoods. Because words from dense neighborhoods have many similar sounding neighbors, those words are hypothesized to contain more phonemic detail; the reverse is true of words from sparse neighborhoods.

According to the lexical restructuring model, words with high neighborhood density should be responded to more accurately than words with low neighborhood density on phoneme awareness tests. This dense word advantage is resultant of more phonemic detail in dense words: words with many neighbors require more phonemic detail to differentiate many similar sounding words. Only a few studies have examined the influence of neighborhood density on phoneme awareness. Metsala (1999) showed that preschool children, ages 3 and 4 years old, were more likely to correctly delete a sound from a word during a phoneme deletion task when the word was from a dense neighborhood. Likewise, a study of phoneme awareness abilities in 5-year olds showed that children were more likely to correctly identify words as rhyming/nonrhyming when the words were from dense neighborhoods (De Cara &

[1]Others have proposed a similar link between vocabulary acquisition and lexical specificity (e.g., Fowler, 1991; Munson et al., 2005). We chose to test the lexical restructuring model because it posits direct hypotheses for the influence of word-level lexical characteristics on phoneme awareness. Not all available evidence supports the model's tenet that specificity increases through vocabulary growth. In particular, studies of toddlers and infants show that early lexical representations are quite detailed (e.g., Bailey & Plunkett, 2002; Swingley, 2003). Task differences may, in part, explain the conflicting evidence: implicit tasks are used to examine representations in young children whereas studies of preschool and school-age children involve tasks that are explicit in nature. Explicit meta-linguistic reflection may require processing influenced by word similarity. Another possibility is that young children may perceive and store sounds differently than older children (cf. Storkel, 2006; Werker & Curtin, 2005). Clearly, further work is needed to determine if lexical restructuring is a product of vocabulary growth, as stated by the model, or if representations reflect specification according to other factors such as processing differences and/or task demands.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Goswami, 2003). This effect has also been shown in adults (Ventura, Kolinsky, Fernandes, Querido, & Morais, 2007).

1.1.2 Word Frequency—

Word frequency refers to the number of times a word is heard in the ambient language. The lexical restructuring model proposes that high frequency words have lexical representations with more phonemic detail which translates into higher phoneme awareness performance on high frequency words compared to low frequency words. The model, however, does not specify the mechanism by which frequency contributes to restructuring. It may be that the predicted word frequency effect is a result of token-based restructuring: a word that is heard produced by multiple speakers, multiple times would contain more phonemic detail as a result of contrasting tokens of phonetic variations of the same word (Richtsmeier, Gerken, Goffman, & Hogan, 2009; vis a vis type-based variation quantified by neighborhood density).

Studies have shown an influence of word frequency on phoneme awareness performance. For example, Troia, Roth, & Yeni-Komshian (1996) found that kindergarten and 2nd grade children performed with greater accuracy on a phoneme awareness task that involved blending sounds in high frequency words compared to those in low frequency words (see also Roth, Troia, Worthington, & Handy, 2006)

1.1.3 Interaction between neighborhood density and word frequency—

The lexical restructuring model (Metsala & Walley, 1998) predicts that both word frequency and neighborhood density should affect phoneme awareness performance. It is also plausible that the two lexical characteristics interact to influence phoneme awareness performance. To date, one study has investigated the influence of both neighborhood density and word frequency on phoneme awareness performance (Garlock, Walley, & Metsala, 2001). Garlock et al. orthogonally crossed the word frequency and neighborhood density of words on two phoneme awareness tasks (i.e., initial phoneme isolation and initial phoneme deletion). To capture potential differences across ages, the task was administered to young children (i.e., preschoolers and kindergartners), older children (i.e., 2nd graders), and adults. Analyses based on a composite measure of performance on both phoneme awareness tasks revealed no statistically significant effects or interactions of neighborhood density or word frequency. Unfortunately potential interactions may have been masked by floor and ceiling effects: the tasks were either too difficult - in the sample of young children - or too easy - in the case of older children and adults (Garlock et al., 2001).

1.2 Study Questions and Predictions

In this study we examined the influence of word frequency and neighborhood density on phoneme deletion performance. Based on past findings, we predicted that phoneme awareness accuracy would increase as the neighborhood density of a test word increased. Likewise, we predicted that phoneme awareness accuracy would increase as the frequency of a test word increased. We also considered an interaction between neighborhood density and word frequency. Although the lexical restructuring model holds no specific predictions, we hypothesized that words will reach a *threshold* of lexical specification due to word frequency or neighborhood density resulting in highly accurate phoneme awareness, such that phoneme awareness performance is unrelated to neighborhood density in words with high frequency and unrelated to word frequency in words with high neighborhood density.

We also add to extant literature by investigating lexical influences on phoneme awareness across development. Based on past findings, we predicted that grade would be positively related to phoneme awareness accuracy. Unique to this study, we predicted that neighborhood density and word frequency would be related to phoneme awareness performance less so in 4th grade compared to 2nd grade. We reasoned that as children's lexicons expand across grades neighborhood density becomes more equivalent across words (see Storkel, 2004a, 2004c) resulting in a reduced relative impact of dense words on phoneme awareness performance. Likewise, word frequency would increase for most words from 2nd to 4th grade reducing the relative impact of words with high frequency on phoneme awareness performance. Taken together we predicted that the interaction between neighborhood density and word frequency on phoneme awareness performance would be less robust in 4th grade compared to 2nd grade.

2. Method

2.1 Participants

Data included phoneme awareness item level performance from a sample of children who completed the same phoneme deletion task in 2nd and 4th grades (n = 194; 113 males, 81 females). These participants were a random subsample of children participating in a longitudinal study of language impairments (n = 570). The sample of 570 children was itself a subsample of over 7000 children who took part in an epidemiologic study of language impairments in kindergarten children (see Tomblin et al., 1997). In terms of demographics, the longitudinal sample, and by extension the random sample, was representative of the region in which the sample was recruited (i.e., state of Iowa - midwest portion of the United States of America).

The phoneme awareness task was administered in kindergarten, 2nd and 4th grades; however, kindergarten data were not examined because very few children completed the phoneme deletion test items. Furthermore, item-level accuracy was not coded when the task was originally scored. For efficiency, we coded item-level accuracy for a random subsample which comprised our 194 participants. Table 1 contains the standard score means and standard deviations for the samples on norm-referenced, standardized measures of receptive vocabulary, nonverbal intelligence, word reading, and reading comprehension as well as raw score phoneme awareness test scores in the 2nd and 4th grades. As shown, our sample of children included a wide range of performance on all measures indicative of the range of performance in typical 2nd and 4th grade classrooms. All children were native English speakers and each resided in homes in which English was the primary language. No child had a differential diagnosis such as Down syndrome or autism upon enrollment in the longitudinal study in kindergarten.

2.2 Phoneme Awareness Test

Each of the 194 participants was administered a phonological awareness deletion task (Catts, Fey, Zhang, & Tomblin, 2001) in 2nd and 4th grades. This task was created to assess phonological awareness abilities during the elementary school years. Although it was not created to examine lexical influences on phoneme awareness, the task contains words varying in both neighborhood density and word frequency, as well as type of deletion (i.e., consonant cluster or singleton) and place of deletion (i.e., deletion in initial consonant(s) versus deletion in final consonant(s)). The task consists of 13 word/syllable deletion test items (e.g., say `cowboy' without `cow') and 17 phoneme deletion test items (e.g., say `fat' without /f/). Of the 17 phoneme deletion items, 8 required the deletion of the first sound in a monosyllabic word, whereas the other 9 items required deletion of the final sounds in a monosyllabic word. Likewise, 7 required deletion of a singleton, whereas 9 required deletion of a consonant within a consonant cluster.

In this study, only the phoneme deletion test items from the task were examined (test items 14–30). Phoneme deletion items were selected for two reasons. First, the tenets of the lexical restructuring model relate most directly to phoneme-level awareness tasks. Neighborhood density is a metric that uses phonemes to determine similarity amongst words in the lexicon; hence, a phoneme-based task would likely show the most sensitivity to performance differences based on a phoneme-level measure such as neighborhood density. Second, a phoneme deletion task seemed a logical choice because it consistently serves as the best predictor of future and concurrent word reading skills when compared to other phoneme-level awareness tasks (Torgesen, Wagner, & Rashotte, 1994).

The deletion task was administered with no basal rule; however a ceiling rule was implemented: if a participant produced six consecutively tested words incorrectly, the test was discontinued. All participants in the random sample completed all test items. That is, each participant had an opportunity to complete all test items in spite of the ceiling rule. Trained speech-language pathologists and/or educators administered the task at each grade as part of a larger test battery. The task did not show evidence of floor or ceiling effects in either the 2nd and 4th grade samples. Including only the phoneme deletion test items, the 2nd grade raw score mean was 9.59 (*SD* = 2.98). In 4th grade the raw score mean was 11.24 (*SD* = 2.50; see Table 1). Using an odd-even split half reliability test, the task evidenced 2nd grade reliability of .86 and 4th grade reliability of .83.

To examine lexical influences on phoneme deletion performance, test words were coded according to neighborhood density and word frequency. Note that performance on 16 of the 17 test words was coded (and included in the analyses); one test word was removed due to its verb tense marking (i.e., test item #28 *fished*). Neighborhood density was computed using a 20,000 word electronic dictionary (Webster's Seventh Collegiate Dictionary, 1967) containing phonemic transcriptions of American English. Test words ranged in neighborhood density from 5 to 32 with a mean of 16.5 (*SD* = 8.76). Word frequency data were also obtained from the electronic dictionary (Kucera & Francis, 1967). Test words ranged in word frequency from 1 to 601 with a mean of 125.81 (*SD* = 179.27). Table 2 contains test words and their corresponding neighborhood density and word frequency counts. Frequency counts were highly skewed; therefore, we transformed them using a natural log transformation. To facilitate interpretation of the results, we converted both neighborhood density and log-transformed word frequency counts to *z* scores based on corresponding counts from 3–5 phoneme words extracted from the full electronic dictionary. Neighborhood density and word frequency *z* score for each test word are listed in Table 2. The mean neighborhood density *z* score was 1.10 (15.87 neighbors), and the mean log word frequency *z* score was 1.07 (equivalent to 41.39 raw frequency), indicating these words were relatively more dense and more frequent than the typical word in the corpus.

Note that neighborhood density and word frequency counts were based on adult data. Previous studies have shown that adult-based data and child-based data produce similar results (e.g., Jusczyk, Luce, & Charles-Luce, 1994). Likewise, adult-based counts have been used in numerous previous studies examining child language (e.g., Newman & German, 2002; Storkel, 2001, 2004b).[2]

2.3 Data Analysis

We employed a multilevel logistic regression predicting a child's response (correct/incorrect) with phoneme awareness test items nested within children and a child-specific intercept to account for individual differences in phoneme awareness. This allowed us to

[2]Child counts were recently made available (Storkel & Hoover, in press). Correlations between our adult counts and the new child counts were high, as expected (word frequency: R = .71; neighborhood density: R = .97). Adult counts were used for analyses because child counts were not available for 3 of our test words. This is likely the result of a smaller child database, which includes 4832 words; the adult database includes over 20,000 words.

determine the influence of lexical factors, neighborhood density and word frequency, on phoneme awareness performance independent of a child's phoneme awareness. We included one child-level variable, grade, expressed as a dummy variable with 0 = grade 2 and 1 = grade 4. Six variables were included as predictors: (1) each word's neighborhood density (z-score), (2) each word's frequency (z-score of the log transformation), (3) an interaction between neighborhood density and word frequency; (4) an interaction between neighborhood density and grade; (5) an interaction between word frequency and grade, and (6) an interaction between neighborhood density, word frequency, and grade. We also included two covariates to control for non-lexical item effects: type of deletion, a dummy variable indicating whether the deleted phoneme was deleted from a consonant cluster[3] (coded as 1) or not (i.e., singleton deletion coded as 0), and place of deletion (deletion in initial consonant or consonant cluster coded as 0, deletion in final consonant or consonant cluster coded as 1), both of which have been shown to affect the difficulty of phoneme deletion (Bruck & Treiman, 1990;Catts, Wilcox, Wood-Jackson, Larrivee, & Scott, 1996). The model was estimated using PROC NLMIXED. The script is available by request.

3. Results

For descriptive purposes, the percentage of children who were able to correctly delete the target phoneme on each test word is presented in Table 2. Results from the multilevel logistic regression are summarized in Table 3. Consistent with standard logistic regression, we discuss results in terms of odds ratios, which describe the relative odds of a correct response.

The lexical restructuring model predicts that words with higher neighborhood density will show more accurate phoneme deletion compared to those with lower neighborhood density. This prediction was supported. For each standard deviation increase in neighborhood density, children were 2.09 more likely to respond correctly ($p < .001$). There was no evidence that the effect of neighborhood density on phoneme deletion changed across grades, as indicated by the nonsignificant interaction of density and grade (odds ratio = .85, $p = .22$).

The lexical restructuring model predicts that words with higher word frequency will show more accurate phoneme deletion compared to those with lower word frequency. This prediction was also supported. For each standard deviation increase in log-transformed word frequency, children were 2.07 times more likely to respond correctly ($p < .001$). The effect of word frequency was moderated by grade, such that the effect was .69 times as large in 4th gradecompared to 2nd grade ($p < .001$), leading to an overall odds ratio of 1.43 for 4th grade (2.07 * .69).

The lexical restructuring model does not hold direct predictions for an interaction between neighborhood density and word frequency; however we expected a threshold effect in which the effect of neighborhood density on phoneme awareness performance would be lower with high word frequency and, conversely, the effect of word frequency would be lower with high neighborhood density. This prediction was supported. A negative interaction was found between neighborhood density and word frequency (odds ratio = 0.45, $p < .001$): as word frequency increases, the effect of neighborhood density on phoneme deletion was reduced. There was also a significant three-way interaction, such that the interaction is smaller in fourth grade than in second grade (odds ratio = 0.45 * 1.35 = 0.61, $p = .02$).

[3]We note that type of deletion is not entirely clear for one word (/fju/; see Davis & Hammond, 1995; Gierut & O'Connor, 2002); however results were not different regardless of the word's `type of deletion' coding.

4. Discussion

This study sought to determine the influence of the two lexical characteristics – neighborhood density and word frequency – on phoneme awareness performance in a longitudinal sample of children tested in 2nd and 4th grades. Of particular interest was a predicted interactive influence of neighborhood density and word frequency on phoneme awareness performance. Further, the influence of these characteristics was examined during development as children progressed from 2nd to 4th grade.

Based on the tenets of the lexical restructuring model coupled with past research findings, it was predicted that words from dense neighborhoods would yield more accurate responses than words from sparse neighborhoods on a phoneme awareness test (De Cara & Goswami, 2003; Metsala, 1999). Likewise, it was predicted that high frequency words would yield higher accuracy than words with low frequency words (Roth et al., 2006). Both predictions were confirmed.

Further, it was predicted that word frequency and neighborhood density would negatively interact to influence phoneme awareness accuracy. This prediction was confirmed: as word frequency increased, the effect of neighborhood density on phoneme deletion was reduced. The reverse was also true. This interaction is interpreted in a framework in which type-based variations (indexed by neighborhood density) or token-based variations (indexed by word frequency) increase word-level phonemic specification which in turn results in word-specific increases in phoneme awareness accuracy. According to this explanation, words with high token-based counts (i.e., high word frequency) do not benefit from phoneme awareness accuracy gains from additional type-based variation (i.e., high neighborhood density). Inversely, adding token-based variation by way of increased word frequency results in limited additional specification of a word's lexical representation in words with high type-based variations (i.e., high neighborhood density). Based on these findings, we propose that the lexical restructuring model should be expanded to include interactive effects of neighborhood density and word frequency: restructuring is the result of either type-based variation – indexed by neighborhood density – or token-based variations – indexed by word frequency. We have labeled this the *threshold effect* in which a word would reach a `threshold' of phoneme awareness accuracy due to type- *or* token-based word-specific lexical specification leaving little room for performance increases from another form of variation (see also Storkel, 2004a for a similar description of threshold effects on word learning).

In addition to finding a significant interaction between neighborhood density and word frequency, developmental differences were also revealed. As expected, phoneme awareness accuracy increased from 2nd grade to 4th grade. Further, we predicted that the effect of lexical characteristics on performance would decrease from 2nd to 4th grade. We reasoned that vocabulary growth would mute the relative difference between words' neighborhood density and word frequency in turn reducing the effect of both on phoneme awareness performance across grades. This prediction was confirmed: the interactive influence of neighborhood density and word frequency on phoneme awareness was present at both grades; however the effect was slightly reduced in 4th grade.

4.1 Clinical Implications

Predictions from the lexical restructuring model, supported by this study, have potential to inform the assessment of phoneme awareness. Recent policy changes have focused on response to intervention (RTI) as a model for diagnosing reading disability in the early grades. Implementation of RTI models requires multiple measures of pre-reading skills at numerous time points as a way to track change over time. Based on the results of this study, systematically varying the neighborhood density and word frequency of words used on tests of phoneme awareness may serve as a way to create meaningful probes: words from dense neighborhoods or with high frequency would be predicted to be the easiest phoneme awareness test items across the grades, while those from sparse neighborhoods or with low word frequency would be the most difficult. It would be predicted that children who exhibit different patterns may be at risk for reading and language learning difficulties. Likewise attention to these lexical characteristics in phoneme awareness treatment (cf. Morrisette & Gierut, 2002) may increase treatment efficacy. These clinical implications are given with caution because the current study did not address these predictions. Clearly future studies are needed.

4.2 Limitations

The lexical restructuring model posits neighborhood density as one metric for quantifying word-level phoneme specificity resultant of vocabulary growth. However other lexical characteristics, such as phonotactic probability (Munson, Kurtz, & Windsor, 2005) and sound-to-letter correspondences (Ventura et al., 2007), are correlated with neighborhood density and have been shown to influence phoneme specificity as well. Future work will need to determine how an array of characteristics associated with individual words interact to influence restructuring and in turn phoneme awareness. Likewise, this study examined these characteristics in one type of phoneme awareness task, phoneme deletion. Lexical influences on phoneme awareness should be examined in a variety of phonological awareness tasks in relation to individual differences in vocabulary acquisition and associated child-specific neighborhood density and word frequency.

This study involved an examination of the influence of lexical characteristics on phoneme awareness performance using item-level data not created for this purpose. However, it is important to make use of rich datasets to take a preliminary look at theoretically driven research questions while acknowledging limitations of post hoc analyses[4]. We included co-variates, place of deletion and type of deletion, to statistically account for influences on phoneme awareness performance not controlled across our lexical characteristics of interest. The predicted interaction between neighborhood density and word frequency was revealed with these controls in place bolstering the validity of the results. However, it is imperative to conduct a controlled experimental study to confirm these findings.

4.3 Conclusions

The purpose of this study was to test the hypothesis that neighborhood density and word frequency would interact to influence phoneme awareness. Findings supported a threshold effect in which words with high neighborhood density *or* high word frequency were more likely to be correct on our phoneme deletion task. This interaction was present in both 2nd and 4th grades although the effect was less robust in 4th grade. Based on these findings we propose that the lexical restructuring model be amended to include these interactions due to token-based variations (i.e., word frequency) or type-based variations (i.e., neighborhood density).

Acknowledgments

This research was supported by NIH Grants awarded to Tiffany Hogan (DC009667), Hugh Catts (DC02726), and Holly Storkel (DC04781, DC08095). The following NIH training grant also supported this research: DC000052.

[4]See volume 41, issue 1 of *Language Speech and Hearing Services in Schools*, a special edition devoted to promoting the use of existing databases to explore theory-driven questions.

The completion of this study was aided considerably by a valuable research team comprising the following: J. Bruce Tomblin, Paula Buckwalter, Marlea O'Brien, Connie Ferguson, Jodi Schwartz, Juanita Limas, Amy Kundel, and Jacqueline Daniels.

References

Adams, MJ. Beginning to read: Thinking and learning about print. MIT Press; Cambridge, MA: 1990.

Anthony JL, Lonigan CJ. The Nature of Phonological Awareness: Converging Evidence From Four Studies of Preschool and Early Grade School Children. Journal of Educational Psychology 2004;96(1):43–55.

Bailey TM, Plunkett K. Phonological specificity in early words. Cognitive Development 2002;17:1265–1282.

Bruck M, Treiman R. Phonological awareness and spelling in normal children and dyslexics: the case of initial consonant clusters. Journal of Experimental Child Psychology 1990;50:156–178. [PubMed: 2398331]

Catts H, Fey M, Zhang X, Tomblin JB. Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. Language, Speech, and Hearing Services in Schools 2001;32:38–50.

Catts, H.; Wilcox, K.; Wood-Jackson, C.; Larrivee, L.; Scott, V. Toward an understanding of phonological awareness. In: Leong, CK.; Joshi, RM., editors. Cross-language studies of learning to read and spell: Phonologic and orthographic processing. Kluwer Academic Publishers; Norwell, MA: 1996.

Davis S, Hammond M. On the status of on-glides in American English. Phonology 1995;12:159–182.

De Cara B, Goswami U. Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. Journal of Child Language 2003;30(3):695–710. [PubMed: 14513474]

Fowler, A. Phonological processes in literacy. Lawrence Erlbaum Associates; 1991. How early phonological development might set the stage for phoneme awareness; p. 97-117.

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;45:468–492.

Gibbs S. Do pictures make a difference? A test of the hypothesis that performance in tests of phonological awareness is eased by the presence of pictures. Educational Psychology in Practice 2003;19(3):219–228.

Gierut JA, O'Connor KM. Precursors to onset clusters in acquisition. Journal of Child Language 2002;29:495–517. [PubMed: 12109361]

Gillon, GT. Phonological awareness: From research to practice. Guilford Press; New York: 2004.

Jusczyk PW, Luce PA, Charles-Luce J. Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language 1994;33:630–645.

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

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear & Hearing 1998;19:1–36. [PubMed: 9504270]

Metsala JL. Young children's phonological awareness and nonword repetition as a function of vocabulary development. Journal of Educational Psychology 1999;91:3–19.

Metsala, JL.; Walley, AC. Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In: Metsala, JL.; Ehri, LC., editors. Word Recognition in Beginning Literacy. Lawrence Erlbaum Associates, Inc.; Mahwah, NJ: 1998. p. 89-120.Publishers

Morrisette ML, Gierut JA. Lexical organization and phonological change in treatment. Journal of Speech, Language, and Hearing Research 2002;45:143–159.

Munson B, Kurtz B, Windsor J. The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repititions of children without specific language impairment. Journal of Speech, Language, and Hearing Research 2005;48:1033–1047.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Newman RS, German DJ. Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. Language and Speech 2002;45(3):285–317. [PubMed: 12693688]

Richtsmeier PT, Gerken L, Goffman L, Hogan T. Statistical frequency in perception affects children's lexical production. Cognition 2009;111:372–377. [PubMed: 19338981]

Roth FP, Troia GA, Worthington CK, Handy D. Promoting awareness of sounds in speech (PASS): The effects of intervention and stimulus characteristics on the blending performance of preschool children with communication impairments. Learning Disability Quarterly 2006;29:67–88.

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

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

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 toddlers. Manuscript submitted for publication. 2006

Storkel HL, Hoover J. An on-line calculator to compute phonotactic probability and neighorhood density based on child corpora of spoken American English. Behavior Research Methods. in press.

Swingley D. Phonetic detail in the developing lexicon. Language and Speech 2003;46:265–294. [PubMed: 14748447]

Tomblin JB, Records N, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. Journal of Speech, Language, and Hearing Research 1997;40:1245–1260.

Torgesen JK, Wagner RK, Rashotte CA. Longitudinal studies of phonological processing and reading. Journal of Learning Disabilities 1994;27(5):276–286. [PubMed: 8006506]

Troia GA, Roth FP, Yeni-Komshian GH. Word frequency and age effects in normally developing children's phonological processing. Journal of Speech and Hearing Disorders 1996;39:1099–1108.

Ventura P, Kolinsky R, Fernandes S, Querido L, Morais J. Lexical restructuring in the absence of literacy. Cognition 2007;105:334–361. [PubMed: 17113063]

Wagner RK, Torgesen JK, Laughon P, Simmons K, Rashotte CA. Development of young readers' phonological processing abilities. Journal of Educational Psychology 1993;85:83–103.

Walley AC, Metsala JL, Garlock VM. Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. Reading and Writing: An Interdisciplinary Journal 2003;16:5–20.

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

Werker JF, Curtin S. PRIMIR: A developmental framework of infant speech processing. Language Learning and Development 2005;1(2):197–234.

Table 1

Standard score means and standard deviations (*n* = 194) on a 2nd grade measure of nonverbal intelligence and 2nd and 4th grade measures of word reading and reading comprehension. Raw score means and standard deviations on phonological awareness deletion task.

	Nonverbal Intelligence[a]	Word Reading[b]	Reading Comprehension[c]	Phoneme Awareness[d]
2nd Grade				
Mean	98.44	100.13	98.79	9.93
SD	14.47	13.19	13.66	2.34
4th Grade				
Mean	Not administered	100.35	98.46	11.65
SD	Not administered	13.32	14.02	2.26

[a]Composite standard score derived from the Picture Completion and Block Design subtests of the *Wechsler Preschool and Primary Scale of Intelligence-Revised* (Wechsler, 1989).

[b]Standard score based on the Word Identification subtest of the *Woodcock Reading Mastery Tests-Revised (WRMT-R*, Woodcock, 1987).

[c]Standard score on Passage Comprehension subtest of the *WRMT-R.*

[d]Raw score out of 16 on the phonological awareness deletion task (Catts et al., 2001).

Table 2

Phonological awareness deletion task item neighborhood density, word frequency, and type of deletion with percentage of children who correctly responded to test words in 2nd and 4th grades

Word[1]	Neighborhood density[2]	Z-score of Neighborhood Density[3]	Frequency[4]	Z-score of Log Frequency[5]	Place of Deletion (initial vs. final)	Type of Deletion (cluster vs. singleton)	Percentage of Children who Correctly Responded in 2nd grade	Percentage of Children who Correctly Responded in 4th grade
f̲æt	28	2.64	60	1.43	initial	singleton	91.7%	95.9%
s̲it	32	3.15	54	1.37	initial	singleton	81.9%	91.3%
ʃ̲aʊ̆t	25	0.99	9	0.37	initial	singleton	88.6%	97.4%
t̲al	27	2.51	55	1.38	initial	singleton	91.7%	97.4%
d̲ɔr	13	0.74	312	2.35	initial	singleton	94.8%	98.5%
f̲ju	12	0.61	601	2.71	initial	cluster	85.0%	86.1%
s̲nel	5	−0.28	1	−0.85	initial	cluster	45.4%	72.3%
θ̲rɛd	6	−0.15	15	0.66	initial	cluster	48.4%	75.4%
tiz̲	16	1.12	6	0.15	final	singleton	82.5%	90.2%
tun̲	27	2.51	10	0.43	final	singleton	86.1%	92.3%
dʌmp̲	13	0.74	312	−0.08	final	cluster	66.0%	79.0%
faɪnd̲	12	0.61	399	2.49	final	cluster	74.7%	86.7%
fæs̲t	15	0.99	78	1.58	final	cluster	24.2%	28.7%
waɪv̲z	5	−0.28	21	0.84	final	cluster	16.5%	16.4%
waɪl̲d	10	0.36	56	1.39	final	cluster	14.9%	25.1%
bɛn̲d	18	1.37	24	0.92	final	cluster	13.4%	24.6%
	M = 16.5 *SD* = 8.76	*M* = 1.10 *SD* = 1.07	*M* = 125.81 *SD* = 179.27	*M* = 1.07 *SD* = 0.97	intial = 8 words	Singleton = 7 words	*M* = 62.90% *SD* = 30.7%	*M* = 72.3.2% *SD* = 30.1%
	Range = 5 – 32	Range = −0.28 – 3.15	Range = 1 – 601	Range = −0.85 – 2.71	final = 8 words	Cluster = 9 words	Range = 13.4% – 94.8%	Range = 16.4% – 98.5%

[1] Words listed in same order as they were presented to child during administration of the task.

[2] Neighborhood density calculation based on a 20,000 word electronic dictionary (Webster's Seventh Collegiate Dictionaryo, 1967) containing phonemic transcriptions of American English.

[3] Neighborhood density *z* scores based on counts from 3–5 phoneme words extracted from the 20,000 word electronic dictionary.

[4] Word frequency obtained using Kucera and Francis (1967) written word frequency counts contained in the 20,000 word electronic dictionary.

[5] Log of word frequency *z* scores based on counts from 3–5 phoneme words extracted from the 20,000 word electronic dictionary.

Table 3

Results of Multilevel Logistic Regression

Predictor	Odds Ratio	95% CI
density	2.09[a]	[1.63, 2.69]
frequency	2.06[b]	[1.81, 2.36]
grade	2.48[c]	[1.89, 3.27]
interaction between density and frequency	0.45	[0.37, 0.55]
interaction between density and grade	0.85	[0.65, 1.10]
interaction between frequency and grade	0.69[c]	[0.58, 0.82]
three-way interaction between density, frequency, and grade	1.35	[1.06, 1.73]
type of deletion	0.12[c]	[0.09, 0.16]
place of deletion	6.14[c]	[4.97, 7.60]

Note. 95% CI is the 95% confidence intervals. Confidence intervals containing 1 are nonsignificant.

[a] odds ratio for z-score.

[b] odds ratio for z-score of natural log transform.

[c] odds ratio for dummy variable **d**