DIFFERENCES BETWEEN EARLY-DEVELOPING AND LATE-DEVELOPING PHONEMES IN PHONOLOGICAL PROCESSING

by

Michelle Wagner Moore

BS, Pennsylvania State University, 1998

MA, Case Western Reserve University, 2002

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Michelle Wagner Moore

It was defended on

April 11, 2012

and approved by

Michael W. Dickey, Ph.D., Assistant Professor
Department of Communication Science and Disorders

Charles A. Perfetti, Ph.D., Distinguished Professor
Department of Psychology

Dissertation Co-Director: Julie A. Fiez, Ph.D., Professor
Departments of Psychology, Neuroscience, and Communication Science and Disorders

Dissertation Director: Connie A. Tompkins, Ph.D., Professor
Department of Communication Science and Disorders

DIFFERENCES BETWEEN EARLY-DEVELOPING AND LATE-DEVELOPING PHONEMES IN PHONOLOGICAL PROCESSING

Michelle W. Moore, Ph.D.

University of Pittsburgh, 2012

Theoretical accounts of communication disorders often hinge on tasks with various confounds. The aim of this study was to challenge the assumption that children with specific language impairment (SLI) have deficits in phonological memory storage capacity solely because they perform poorly on nonword repetition tasks. This assumption was tested using a novel contrast of early- and late-developing phonemes that was predicted to elicit differences in nonword repetition performance even after controlling for confounding factors. Using a differential diagnosis model of testing, a variety of tasks were administered to determine if early vs. late phoneme differences (ELP) would persist after auditory perceptual, articulatory, phonological memory storage capacity, and lexical demands were minimized. In Study 1, 30 undergraduates completed nonword repetition, nonword reading, and auditory lexical decision tasks in which half of the stimuli contained only early-developing phonemes and half contained only later-developing phonemes. In Study 2, the ELP contrast was examined in another group of 20 undergraduates who completed auditory and visual lexical decision with and without concurrent articulation.

Both nonword accuracy and word-nonword discriminability were consistently lower for items with later-developing phonemes than for those with early-developing phonemes, but there were no differences in response times. Results support the growing literature suggesting that nonword repetition relies on multiple processes and cannot be used as a measure of phonological

memory storage capacity alone. Additionally, nonword repetition performance draws on skills apart from auditory perceptual demands, articulatory demands, and lexical knowledge. This in turn challenges the assumption that children with SLI have deficits in phonological memory storage capacity simply because they perform poorly on nonword repetition.

The results may suggest that the ELP contrast reflects differences in the quality of the phonological representations that derive from the timing of phoneme acquisition, though other possible explanations for the differences are discussed (e.g., other articulatory influences). The ELP manipulation within this battery of tasks affords many possible outcomes that might adjudicate between the possible accounts of deficits that have been associated with SLI, such as perceptual and motor speech deficits that may each contribute independently or additively to poor performance in phonological processing.

TABLE OF CONTENTS

LIST OF TABLES

PREFACE

I am beyond grateful to have had an amazing network of people who have impacted my life in immeasurable ways. First, I thank those who I have met through the academic community. As I've said before, there is no "I" in research. To friends and colleagues in the Fiez lab, the Tompkins lab, the LRDC, and the Department of CSD, you helped with some of the nuts and bolts of this project and provided insight, scientific discussion, encouragement, and an empathetic perspective of the highs and lows of the research process. Thank you for riding this roller coaster with me.

To mentors like Chris Dollaghan and Melanie Schuele who took me on as a student in earlier stages of my academic career, your obligation to advise me has ended, but your willingness to continue to offer guidance speaks volumes of your character and your mentorship. Thank you for sharing your perspectives and for being committed to helping me. I have so greatly learned from you.

To my committee – Connie, Julie, Mike and Chuck, often times I have said that I had the ideal dissertation committee. The value of your knowledge and experience goes without saying, but your legacy extends far beyond the margins of your CVs. You are my science coaches, my role models, and my friends. Thank you.

Julie, it's been a privilege to work with you and to be a part of your lab. Your research and passion for science are inspiring, and you have provided me with such a rich and diverse

environment for learning and growing as a scientist. Your mentoring has been instrumental in helping me to succeed in my training, overcome challenges, and take the next steps in my career. I feel so fortunate to have trained under your wing. Thank you.

Connie, you have pushed me to grow as a critical thinker and researcher, providing invaluable mentoring and encouragement when aspects of the process became particularly challenging. Beyond simply advising me through stages of the program requirements, you have sacrificed time and efforts to help prepare me for a career as a researcher and to share bits of life. I can't thank you enough for investing in me.

To my friends and family, you are my village. Every one of you has influenced my life in so many ways that I couldn't begin to capture it all here. You have given me laughter, love, and joy. You have celebrated with me during victories and shared my burdens during trials. You have been my greatest advocates and my teachers. This is not about how you have influenced me during my doctoral studies; this is about how you have blessed my entire life. You are my roots. Thank you.

Mama, I remember one Christmas morning during my childhood when one of my gifts from you was stationery. You explained to another family member that I liked to write. I didn't realize that about myself at the time, but you were right. You saw my talents and interests before I did and you nurtured them. And now I have written a dissertation. You are a shining light in my life and the real gift to me is that I am your daughter.

Brian, you are amazing. You learned what a phoneme was so that you could be an active listener when I talked about my research. You have been steadfast when I have struggled and picked up my slack when my work became consuming. There is no question that I am a stronger and better person because you are by my side. Thank you for brightening my days. I am so grateful for you.

"Now to Him who is able to do immeasurably more than all we ask or imagine, according to His power that is at work within us, to Him be glory." Ephesians 3: 20, 21.

1.0 INTRODUCTION

The phonological processing skills of children with communication disorders have been widely studied. Of particular interest has been the well-documented finding that many children with certain communication disorders, including specific language impairment (SLI) and reading disabilities, perform poorly on nonword repetition tasks compared to their typically-developing peers. There is a rich literature demonstrating poor performance on nonword repetition in samples of children with SLI (for a review, Graf-Estes, Evans, & Else-Quest, 2007), but there are contradictions in how the task is used and interpreted, and how it can inform the different accounts of SLI. Although it has been well-supported that the ability to repeat nonwords is a skill that relies on multiple processes (Archibald & Gathercole, 2006; Briscoe, Bishop, & Norbury, 2001; Coady & Evans, 2008; Edwards & Lahey, 1998; Gathercole, 2006; Graf-Estes, et al., 2007; Gupta, 2006; Snowling, Chiat, & Hulme, 1991), nonword repetition has been used broadly as a classic measure of phonological memory storage capacity, thus creating an inferential error in interpreting the deficits associated with SLI (i.e., children with SLI have phonological memory storage capacity deficits because they perform poorly on NWR).

In order to better understand the underlying processes involved in NWR as well as the deficits associated with language impairment, previous work has primarily explored the contribution of lexical knowledge, memory demands, and articulatory complexity in nonword repetition performance. Researchers agree that repeating a nonword also requires "robust representations of the underlying speech units" (Coady & Evans, 2008, p. 25), but it is unclear what role underlying phonological representations – independent from memory and other lexical factors – play in NWR performance because of the difficulty in conceptualizing and measuring them.

The aim of this study was to challenge the assumption that children with SLI have deficits in phonological memory storage capacity because they perform poorly on nonword repetition tasks. This assumption was tested using a novel contrast of early- and late-developing phonemes that were predicted to elicit differences in nonword repetition performance even after many potential confounding factors were controlled. If decreased performance in nonword repetition can occur apart from other factors (e.g., phonological memory storage capacity and lexical knowledge), then this suggests that the locus of the phonological deficit in SLI may not necessarily be the result of phonological memory storage capacity deficits.

2.0 BACKGROUND

2.1 PHONOLOGICAL PROCESSING AND SLI

The construct of phonological processing and other related concepts (e.g., phonological segmentation, phonological encoding, phonological memory, and so on) have often been used loosely, inconsistently, and/or without a clear definition in the literature. For example, in some instances phonological processing is described as a group of skills or mechanisms that includes phonological awareness and phonological memory (e.g., Wagner, Torgesen, & Rashotte, 1999), whereas other bodies of work distinguish these three as separate processes (e.g., Gathercole, 2006). This confusion is at least in part due to the abstractness of these phonological concepts as well as the interactive nature of the different processes that are involved in speech and language (e.g., Dell, 1986; Levelt, Roelofs, & Meyer, 1999). This current discussion will work from a general definition for phonological processing put forth by Gupta and MacWhinney (1997): "Phonological processes involve knowledge of the sound pattern of a particular language and use of this knowledge in the process of identifying or producing spoken words" (p.306; see also a similar definition by Wagner & Torgesen (1987) that extends this idea to both oral and written words).

'Phonological representation' is another concept that has been increasingly studied in the SLI literature, but its abstract nature and the methodological variations used to examine this construct also are confusing (Anthony, et al., 2011). A general, well-accepted definition for a phonological representation is that it is the long-term memory store of a word's phonological information (Stackhouse & Wells, 1997; Sutherland & Gillon, 2005). Consistent with this definition but elaborating further, Hester and Hodson (2004) state:

> "The term 'representation' is used to describe how information is held in a system. In the brain, representations are seen as cortical patterns of synaptic connectivity (Elman, et al., 1996). Individuals code linguistic information in abstract mental representations of the various subsystems of language, including phonological, semantic and syntactic forms. Mental representations at the phonological level comprise codes for the individual sounds of a language and rules for ordering and combining them (Wolf, Vellutino, & Berko Gleason, 1998, p. 115)."

This current project works from the definitions provided above. A phonological representation is an abstract construct used to describe a neural network that stores phonological information. Phonological representations are just one type of mental representation used when processing lexical information. Hester and Hodson (2004) also mention semantic and syntactic representations, and I add that there are orthographic representations (the visual form, or spelling, of words) to consider in lexical processing as well (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989). As mentioned previously, it is also important to note that these mental representations do not act in isolation, but rather are interactive during lexical processing (e.g., Perfetti & Hart, 2002; Taft, 2006).

Specific language impairment is a label used to describe children who have an impaired language system in an otherwise seemingly intact cognitive system. Certainly there is a broad range of deficits related to the form, content, and use of language that can be associated with language impairment, but two prominent lines of work in SLI have focused on grammar and phonology. The focus of the current research is on the phonological processing accounts of SLI (discussed below), but first an example of the interconnections between grammar and phonology is briefly mentioned.

One example that has been used to demonstrate the potential influence of phonology on morphology is past tense –ed, a common morphological marker often omitted in the language of children with SLI (Joanisse & Seidenberg, 1998). In English, there are three allophonic variations for past tense –ed: 1) /t/ is used when the final phoneme of the verb is a voiceless consonant (e.g., jumped is /dʒʌmpt/); 2) /d/ is used when the final phoneme of the verb is a voiced consonant or vowel (e.g., played is /pled/); 3) /ɪd/ is used when the final consonant of the verb is an alveolar /t/ or /d/ (e.g., weeded is /widɪd/). Each of these uses of past tense –ed requires a phonological decision, namely which allophonic variation should be used based on the phonological characteristics of the final phoneme in the verb. Thus, this example demonstrates how a phonological deficit can play a causal role in the ostensive impairment of grammatical morphemes and syntax (see also Chiat, 2001). In fact, deficits in phonological processing are thought to impact a multitude of verbal functions, including speech production (Sutherland & Gillon, 2007), phonological awareness and decoding during reading (Elbro, 1996; Elbro, Borstrom, & Petersen, 1998; Sutherland & Gillon, 2007), nonword repetition (e.g., Gathercole, 2006), and other language skills (Joanisse & Seidenberg, 1998). The potential impact is great, affecting multiple areas of communication.

Within the umbrella of a phonological deficit hypothesis in SLI, several accounts provide more specific claims about the locus of impairment in phonological processing. Three general hypotheses are that the phonological deficits in SLI are due to 1) a deficit in phonological storage capacity, 2) underspecified phonological representations due to prolonged changes in lexical representations and knowledge, or 3) underspecified phonological representations not mediated by changes in lexical knowledge (i.e., core phonological processing).

According to the phonological storage capacity account (PSCA), children with language impairment have difficulty processing phonological information because of the limited capacity of their online phonological store (i.e., the "set of currently activated phonological representations"; Gathercole, 2006, p. 522; Gathercole & Martin, 1996). With a limited capacity to maintain phonological information in short-term memory, less online information is available during the phonological processes involved in verbal functions such as word-learning, sentence formulation, and so on, thus compromising performance. The related body of literature is not always clear in distinguishing phonological storage capacity from phonological memory; these concepts are often used interchangeably, and deficits in phonological storage capacity have also been referred to as deficits in phonological memory (Gathercole, 2006; Gathercole & Baddeley, 1990; Graf-Estes, et al., 2007; G. Jones, Tamburelli, Watson, Gobet, & Pine, 2010). However, it should be noted that the two concepts are not synonymous. Phonological storage capacity is the maximum quantity, or span of phonological representations that can be activated without compromising the quality of the representations. Phonological memory is defined more generally as the ability to temporarily maintain the sounds in a language for short-term manipulation and retrieval. The relationship is not bidirectional – although phonological storage capacity deficits will usually affect phonological memory, phonological memory deficits cannot always be attributed to a limited storage capacity. For example, if the quality of the phonological representations is underspecified or degraded, this could be a direct cause for poorer performance (e.g., because the incoming information is not accurately coded) or it might indirectly affect performance by making it harder to maintain the information (e.g., because the activation is weaker and thus more prone to decay; Bowey, 2006; Edwards & Lahey, 1998).

The second general hypothesis to explain the phonological deficits in SLI proposes that children with SLI have underspecified phonological representations due to prolonged changes in lexical representations and knowledge. This hypothesis stems from accounts of vocabulary acquisition (e.g., lexical restructuring hypothesis, Bowey, 2001; Bowey, 2006; lexical reorganization account, Metsala & Chisholm, 2010; see also, Munson, Edwards, & Beckman, 2005; Walley, 1993). Metsala and Chisholm (2010) describe a basic principle in their summary of the lexical reorganization hypothesis: "According to this position, increases in vocabulary knowledge impacts [sic] the detail, structure, connection strengths, or autonomy of representations in long-term lexical memory, which in turn, facilitates nonword repetition and vocabulary acquisition" (p.490). According to this view, deficits or delays in vocabulary acquisition could interrupt the development of detailed information in long-term lexical memory such that underspecified information at the syllable or phonological level would be available (e.g., protracted lexical restructuring hypothesis; Bowey, 2006).

The third hypothesis suggests that children with SLI could have deficits in phonological processing that are not mediated by lexical knowledge due to the quality of phonological representations and the ability to retrieve them from long-term memory. Phonological representations can be poor in quality, i.e. underspecified or degraded, due to any number of reasons – for example, perceptual deficits, production deficits, or limited exposure. Whatever the reason, these poor-quality phonological representations are thought to be more prone to decay and interference during lexical processing or other phonological processing tasks, thus compromising performance on the tasks (Anthony, et al., 2011; Bowey, 2006; Coady & Evans, 2008; Edwards & Lahey, 1998; Joanisse & Seidenberg, 1998). Because this hypothesis proposes that the deficit in phonological processing is not mediated by other functions such as lexical knowledge, word learning, or phonological storage capacity, it will be referred to here as a core phonological processing deficit hypothesis.

To summarize, although the classic deficits associated with specific language impairment are related to grammar and morphology, there is an expansive body of work examining the role of impaired phonological processing in SLI. Three general hypotheses have been put forth to provide more specific claims about the locus of phonological processing deficits. The first account described above, the PSCA, proposes that children with SLI have a deficit in phonological memory storage capacity. According to the second account, when lexical acquisition is delayed or impaired, the quality of the phonological representations is underspecified or degraded, thus negatively impacting performance during online phonological processing. A third account, the core phonological processing account, suggests that the quality of the long-term phonological representations and the underlying mechanisms involved in phonological processing in children with SLI can be impaired independently from lexical knowledge and phonological storage capacity.

2.2 A MEASURE FOR PHONOLOGICAL MEMORY STORAGE CAPACITY?

Because nonword repetition has been used so often as a measure for phonological memory storage capacity, and because many children with SLI perform poorly on nonword repetition tasks, a conclusion that is frequently made in the literature is that children with SLI have phonological memory storage capacity deficits. This, however, is an invalid conclusion based on a logical reasoning error:

> Phonological memory storage capacity deficits lead to poor performance on NWR tasks.
> Children with SLI perform poorly on NWR tasks.
>> Therefore all children with SLI have phonological memory storage capacity deficits.

Although both of the assertions may be true, the conclusion is invalid. This is made clear in a more concrete, but similarly-structured example:

> All females are mortal.
> All children are mortal.
>> Therefore all children are females.

In this concrete example, it is clear that, although both assertions are true, the conclusion that all children are females is not true. *Some* children are females, but not all of them. Likewise, *some* children with SLI may have deficits in phonological memory as a result of a limited storage capacity according to the PSCA, but based on NWR performance alone we cannot conclude that *all* children with SLI have a limited phonological storage capacity (e.g., Archibald & Gathercole, 2006; Coady & Evans, 2008; Graf-Estes, et al., 2007).

In fact, there is evidence to suggest that not all poor NWR performance is linked to phonological memory storage capacity deficits. In a study of children with SLI, Archibald & Gathercole (2005; see Gathercole, 2006) administered a nonword repetition task as well as a serial recall task – a classic working memory task used in the cognitive psychology literature where the idea of a phonological store was promoted (e.g., Baddeley, Gathercole, & Papagno, 1998; D. M. Jones, Macken, & Nicholls, 2004), in which a list of items is presented one at a time and the participant is asked to recall the list in the correct order. Compared to their typically-developing peers, the magnitude of the performance difference between the two groups of children on the serial recall task was smaller than on the nonword repetition task, suggesting that another factor was contributing to the NWR outcome.

Perhaps a better measure of phonological memory storage capacity is the manipulation of nonword length within the nonword repetition task. While the NWR task does rely on multiple processes, with this manipulation the focus is on the effect of length – i.e., the difference between shorter and longer items when all other factors are held constant. A phonological memory storage capacity deficit has been demonstrated in children with SLI by their increased sensitivity to the word length effect during nonword repetition tasks – i.e., as the nonword stimuli become longer, children with SLI have an increasingly more difficult time repeating the nonwords as compared to typically-developing children (e.g., Dollaghan & Campbell, 1998). However, in some studies the children with SLI performed more poorly compared to typically-developing controls even on one-syllable nonwords where the capacity demands are not as great, further suggesting that phonological memory storage capacity deficits cannot entirely explain the deficits associated with SLI (for a review, Graf-Estes, et al., 2007). As mentioned previously, it is also reasonable to explain the deficits in phonological memory storage capacity in terms of the quality of the phonological representations. If the quality of the representations is underspecified or degraded, then memory performance would be compromised, particularly in instances where the system is taxed (i.e., presented with longer stimuli; Bowey, 2006).

Other components of nonword repetition have been examined as well, like the articulatory complexity of the nonword stimuli (speech production influences) and the wordlikeness of the stimuli (lexical influences from long-term memory). To examine the contribution of articulatory demands in nonword repetition performance, investigators have compared consonant classes as well as consonant clusters versus consonant singletons in nonwords. Results have generally showed decreased repetition accuracy on more complex stimuli (e.g. nonwords containing consonant clusters) compared to less complex stimuli (e.g. nonwords containing consonant singletons); however, group differences between children with SLI and non-impaired controls have not always been found (Archibald & Gathercole, 2006; Bishop, North, & Donlan, 1996; Briscoe, et al., 2001; Edwards & Lahey, 1998; Gathercole & Baddeley, 1990).

The use of nonwords was once thought to be a way to circumvent lexical influence, but work since then has demonstrated a performance advantage for nonwords that are considered to be more word-like (i.e., more closely resemble a real word) compared to nonwords that are less word-like, suggesting that there can in fact be an influence from long-term lexical knowledge in nonword repetition (Gathercole, 1995; Graf-Estes, et al., 2007). Further support for this conclusion comes from studies in which non-impaired language groups have improved performance when the stressed syllables of the nonwords are real words (Dollaghan, Biber, & Campbell, 1993, 1995), when the neighborhood density is greater (Metsala & Chisholm, 2010), and when the nonwords have a higher phonotactic probability (i.e., the frequency of occurrence of a phoneme or group of phonemes within a given word position in a language; e.g., Edwards, Beckman, & Munson, 2004; Munson, Kurtz, & Windsor, 2005). Taken together, there is substantial evidence to suggest that long-term lexical knowledge can influence performance in nonword repetition.

However, it is not well understood how long-term lexical knowledge contributes to the nonword repetition deficits in children with SLI. In some previous work, where wordlikeness was based on adult ratings of the nonwords, there were no significant differences in the magnitude of the wordlikeness effect when comparing children with SLI and the normal-language control groups (Archibald & Gathercole, 2006; Briscoe, et al., 2001). In contrast to these results, Munson and colleagues (2005) did find differences in performance between impaired and non-impaired groups when manipulating phonotactic probability. They further noted that wordlikeness judgments were not as strong a predictor as phonotactic probability of nonword repetition accuracy.

Could deficits in nonword repetition in children with SLI be explained by more pervasive deficits in processing phonological representations? Although the wordlikeness manipulation and other measures of lexical knowledge (e.g., neighborhood density and phonotactic probability) contribute somewhat to our understanding of the influence of long-term phonological knowledge on nonword repetition, these manipulations are confounded with other aspects of word knowledge that make it difficult to draw any solid conclusions about the role of core phonological processing on performance. For example, in a contrast between a nonword that has no resemblance to an English word versus a nonword that is very similar to a real English word, the latter not only draws upon lexical phonological representations from long-term memory, but potentially could have added influence from semantic and orthographic knowledge as well.

This is a confound that is present in several other studies of SLI that have aimed to examine the underlying phonological representations using other tasks, for example naming and gating (presenting increasingly longer segments of a word until the word is recognized). The inclusion of lexical items in both of these tasks potentially draws upon other aspects of word knowledge so that it is difficult to make conclusions about core phonological processing function. Furthermore, naming, gating, and other similar tasks rely on multiple processes, a factor that seems to be well accepted, but not always well accounted for when assessing phonological processing in child language impairment. Additionally, although a task may in fact elicit the retrieval and processing of phonological information from long-term memory, it also almost always has more peripheral demands as well, such as some sort of perceptual input (receptive tasks) or speech motor output (production tasks).

The purpose of this study was to challenge the assumption that performance differences in nonword repetition necessarily reflect differences in phonological memory storage capacity by using a manipulation of early-developing and late-developing phonemes. If performance differences can be observed after minimizing phonological storage requirements and other factors known to influence nonword repetition performance (e.g., lexical factors), then this would provide evidence that performance on the task necessarily cannot be taken as evidence of a specific problem in phonological storage capacity. This in turn would challenge the current phonological processing accounts of SLI.

2.3 EARLY- AND LATE-DEVELOPING PHONEMES

Moore, Tompkins, and Dollaghan (2010) conducted a study with young adults to examine the psychometric properties of a new nonword repetition task. The nonwords in the task contained only later-developing phonemes, designed with the goal of increasing the articulatory demands and thus the performance demands of the nonword repetition task. As predicted, participants performed more poorly on this task than on a nonword repetition task containing only earlier developing phonemes.

The following discussion focuses on the selection of early- and late-developing phonemes and potential causes for early-late phoneme differences. Regardless of cause, one secondary consequence may be that there are differences in the core phonological representations of early versus late phonemes. If true, this would have important implications for our understanding of early-late phoneme differences in general, as well as their potential for revealing differences in nonword repetition performance.

2.3.1 Selection of phoneme groups.

The phonemes in the Nonword Repetition Task (NRT; Dollaghan & Campbell, 1998) and the Late-8 Nonword Repetition Task (L8NRT; Moore, et al., 2010) were selected based on the consonant mastery framework of Shriberg and colleagues (Shriberg, 1993; Shriberg & Kwiatkowski, 1994). Shriberg (1993) identified the profile of consonant mastery for 64 children with speech delay based on analysis of spontaneous conversational speech samples. The Percentage of Consonants Correct (PCC) for the 24 English consonants was charted in decreasing order, such that natural groupings of 8 were formed – *Early-8* (averaging over 75% correct: /m, b, j, n, w, d, p, h/), *Middle-8* (averaging 25-75% correct: /t, ŋ, k, g, f, v, tʃ, dʒ/), and *Late-8* (averaging less than 25% correct: /s, z, l, r, ʃ, ʒ, θ, ð/).

In the same study, Shriberg compared this trajectory of consonant mastery with six normative studies to determine if the developmental order of acquisition was similar. In the normative studies, anywhere from 54 – 71% of the 24 consonants fell into the expected Early-Middle-Late group, a fairly reasonable but imperfect index of validity for the consonant groups. However, the considerable methodological differences between Shriberg's study and the six normative studies could account for the discrepancy between studies. To name a few, there were methodological differences in criterion level for consonant mastery, age, phoneme sampling methods and transcription. For example, only one of the studies used continuous speech samples to evaluate age of consonant acquisition.

To further validate the Early-Middle-Late framework, Shriberg and Kwiatkowski (1994) compared the PCC profiles of 64 3- to 6-year-old children with speech delay to the 72 3- to 6-year-old children with typically developing speech described in Hoffmann (1982), Hoffmann and Shriberg (1982), and Shriberg (1986, 1993). Shriberg and Kiatkowski noted that all speech-sampling, transcription, and data reduction procedures for the children with normal speech were the same as those used with the speech-delayed children (Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986). Shriberg and Kiatkowski's comparison of PCC profiles showed that, although the children with speech delay on average have decreased PCCs, both groups of children have a similar order of acquisition. For the group of typical-speech-developing children, the average percentages of consonants correct for the Early-8, Middle-8, and Late-8 consonants were 98%, 93%, and 42%, respectively. Thus, although there may be some variation for the order of acquisition for individual consonants depending on the methods used to obtain consonant mastery data, these results provide compelling evidence of the relatively late mastery of the Late-8 consonants (Shriberg, 1993).

Based on this work, Dollaghan and Campbell (1998) designed the NRT using only phonemes from the Early-8 and Middle-8 consonant groups. As suggested by its name, the L8NRT is comprised only of consonants from the Late-8 group (Moore, et al., 2010).

2.3.2 Potential factors contributing to phoneme use.

There has been debate in the literature about the factors that influence the time course at which different phonemes emerge, and that affect performance on tasks that contrast early versus late phonemes. One factor that might be considered is the frequency of use of these sounds in English. From the conversational samples of both children and adults, however, there is no evidence to suggest that the members of one of these phoneme groups occur more frequently (Mader, 1954; Mines, Hanson, & Shoup, 1978). Other more likely potential factors are the perceptual salience and articulatory demands of early versus late phonemes, reviewed below.

2.3.2.1 Perceptual salience of phonemes.

When hearing ability is normal and environmental noise/distortions are minimized, do the variable qualities of phoneme salience impact processing performance in tasks like nonword repetition? It is reasonable to postulate that less audible or more confusable phonemes could be more difficult for infants to learn and, relative to sounds that are very easy to perceive, could also be difficult to discern for a skilled adult listener. On one hand, the infant perception literature convincingly shows that infants are able to perceive even subtle phonetic contrasts, for example differences between native and non-native phonemes, phoneme categories, prosodic features, and so on (for a review, Jusczyk, 1992). However, so often in oral communication, visual cues and verbal contexts are available to assist in auditory perception. Without these aids, like in a nonword repetition task where stimuli are only presented aurally (i.e. no visual cues) and consist of nonwords (i.e. minimized verbal context), we can begin to examine the perceptual salience of phonemes.

Although Moore et al. (2010) primarily attributed increased difficulty with later-developing phonemes to an increase in articulatory demands, in their discussion they noted other possible factors that could influence performance with these later-learned phonemes, like the perceptual salience of the phonemes within the nonwords. In their error analysis, Moore et al. determined that 50% of the errors on the nonword repetition task containing only Late-8 consonants (the Late-8 Nonword Repetition Task, or L8NRT) were substitutions or deletions of the 'th' sounds (/θ, ð/). They discussed previous work that suggests these sounds are particularly difficult to discern and can be easily confused with other phonemes (/f/-/θ/, /v/-/ð/) without the presence of any visual cues or verbal context (Miller & Nicely, 1955). Nevertheless, results from this study alone are still inconclusive because it is difficult to determine if these errors were due to the decreased perceptual salience of the phonemes /θ/ and /ð/, or if other factors like increased articulatory demands played a more dominant role in performance outcomes.

Additional evidence for variation in the perceptual salience of phonemes comes from a study by Redford and Diehl (1999). They examined the perceptual confusability of a set of highly frequent consonants (/p, t, k, f, θ, s, ʃ/). CVC target words were designed using these consonants and embedded in carrier sentences. Participants were asked to listen to the sentences through earphones (in the presence of low-level pink noise, 15-dB signal-to-noise ratio) and write down the target word. Results showed that non-strident (i.e. non-sibilant) fricatives (/f, θ/) had higher percents of error than both the strident (i.e. sibilant) fricatives (/s, ʃ/) and the stop consonants (/p, t, k/). The /ʃ/ had the fewest errors of any of the consonants; the /θ/ had the most errors.

Although Redford and Diehl's (1999) work does provide evidence for variation in phoneme perceptibility, their results suggest that perceptual salience cannot entirely explain the decreased performance on later-developing phonemes, because not all of the late phonemes used in their study (/θ, s, ʃ/) had high confusability. To the contrary, one late phoneme had very high confusability (/θ/) yet another (/ʃ/) had the fewest confusability errors of all consonants measured.

This previous work suggests that decreased perceptual salience may contribute to decreased performance for some – but not all – later-developing phonemes (Moore, et al., 2010; Redford & Diehl, 1999). There are also some early-developing phonemes (e.g. /f, v/) that are susceptible to confusability as well, further suggesting that the perceptual salience of the phonemes is not the sole factor contributing to differences in performance between early- and late-developing phonemes.

2.3.2.2 Articulatory demands of phonemes.

Another consideration for what could contribute to decreased performance with late-developing phonemes is their articulatory demands. When motor ability is normal and environmental impediments are absent, do the variable articulatory demands of phonemes impact processing performance in language tasks?

The speech system is a complex motor system that develops over time. Children are able to produce increasingly more complex sounds with increasing accuracy as, among other developments, their motor system matures and becomes more precise in making fine-grained motor movements. It is reasonable to postulate that phonemes high in articulatory demands are not only difficult to acquire, but will always be relatively more difficult to produce, even for skilled speakers with decades of practice. Just as the coordination of hand and vision to grab an object is a more complex task with more fine-grain tuning than simply moving a hand up and down, some phonemes may have more complexity involved in production.

Kent (1992) inferred the increasing motor demands in phonological development. Based on the order of phoneme mastery described in Sander (1972), Kent considered four developmental sets or motoric stages of sound acquisition, displayed in Table 1.

Table 1. Stages of phonological development based on increased motor demands taken directly from Kent (1992).

Set 1	-Rapid ("ballistic") articulatory movement: /p, m, n/ -Slow ("ramp" movements characterized by constant velocity over a relatively long duration) articulatory movement: /w, h/ -Velopharyngeal valving: stops and nasals present -Voicing adjustment: voiced and voiceless items present -Primary places of articulation: bilabial, alveolar, and glottal	/p, m, n, w, h/
Set 2	-Additional items in the rapid or ballistic movement category: /b, k, g, d/ -Additional items in the ramp movement category: /j/ -Fine force regulation for frication: /f/ -Additional primary place of articulation: velar	/b, d, g, k, j, f/
Set 3	-Additional items in the rapid or ballistic movement category: /t, ŋ/ -Velopharyngeal valving to distinguish /m/-/b, p/ and /n/-/d, t/ -Voicing adjustment to distinguish /p/-/b/, /t/-/d/, and /k/-/g/ -Tongue configuration (bending) for /r/ and /l/	/t, ŋ, r, l/
Set 4	-Tongue configuration for dental, alveolar, and palatal fricatives -Fine force regulation for frication at each place of fricative articulation	/s, z, ʃ, v, θ, ð, ʒ, tʃ, dʒ/

The general order of acquisition depicted in Table 1 is similar to the Early-Middle-Late groupings described earlier (Shriberg & Kwiatkowski, 1994) and thus suggests that there may be a component of articulatory complexity (based on increasing motoric adjustments) to the Early-8, Middle-8, and Late-8 groups. The work of Stokes and Surendran (2005) further supports the role of articulatory complexity in phoneme acquisition. They considered frequency of occurrence, functional load (effects from losing or merging a minimal contrast), and articulatory complexity in phoneme acquisition. The influence of these factors varied by language, but for English, the best predictor for accuracy of consonant production in two-year-olds was articulatory complexity, which accounted for 40% of the variance in their study.

Another way to examine the articulatory demands of phonemes is with diadochokinetic (DDK) rates. DDK tasks are a common diagnostic tool in speech-language pathology used to assess the articulatory rate capacities of children and adults being evaluated for potential motor speech disorders. Patients are asked to repeat single- or multi-syllable phrases (typically /pʌ/, /tʌ/, /kʌ/, /pʌtʌkʌ/) as clearly and quickly as they can. Performance is measured either by number of syllables spoken in a given time, or by amount of time needed to repeat a set number of syllables (Kent, Kent, & Rosenbek, 1987). The theoretical bases of the diagnostic utility of DDK rates has been questioned (Tiffany, 1980), but this is a useful measure for the purpose of understanding the varied articulatory demands of phonemes. Namely, if Late-8 phonemes are more difficult to produce, then one would expect that DDK rates (syllables/second) for these phonemes would be slower compared to Early-8 phonemes that are less demanding.

Blumberger, Sullivan, and Clément (1995) conducted a study in which DDK rates using a broad range of consonant phonemes were examined in adults with traumatic brain injury (TBI) and non-brain-injured controls. Table 2 lists the syllables as well as the data for the control subjects (the data for the patients with TBI are not included since the primary interest is the effect of stimuli variation within the context of typical motor speech abilities). Fewer than half of the Late-8 phonemes were tested, but from the data available there are no apparent differences in DDK rates between the Early-, Middle-, and Late-8 groups in a sample of adults, an unexpected finding given the literature reviewed thus far.

Table 2. DDK rates for normal adults (n = 28). Table adapted from Blumberger, Sullivan, & Clément, 1995.

Early-8	***M*** *	***SD***	**Middle-8**	***M*** *	***SD***	**Late-8**	***M*** *	***SD***
BA	5.78	1.28	TA	6.00	1.07	SA	4.32	1.32
YE	3.19	1.25	KA	5.45	1.06	TH	4.86	1.21
YA	3.31	1.02	GA	4.97	1.08	THA	4.37	1.39
HA	4.18	1.53	FA	5.14	1.34			
DA	5.59	1.21	VA	4.30	1.24			
PA	6.34	1.13	CH	5.09	1.18			
			JA	4.76	1.23			

* - Scores reported are the average number of syllables per second
Note. Stimuli are listed here as they were presented in the original work.

Articulatory complexity has been addressed in studies using language processing tasks as well. In examining the role of articulatory complexity in nonword repetition performance in children, for example, many studies have compared performance on nonwords with consonant clusters versus nonwords with only consonant singletons (e.g., Archibald & Gathercole, 2006; Bishop, et al., 1996). Although these descriptive studies begin to address the factor of articulatory demands, it is confounded with factors such as a potential mismatch in number of phonemes and duration of stimuli, with nonwords containing consonant clusters potentially having more phonemes and longer durations (Gathercole, 2006). A possible way to control for these potential confounds is to consider other factors that would increase the articulatory demands, such as increasing motor demands required to produce different phonemes.

Edwards and Lahey (1998) attempted to do this by exploring whether the increased articulatory demands required by fricative and liquid consonant phonemes – suggested by Kent (1992) to require increased motor control – affected repetition of nonwords in children with SLI. Results were mixed with performance on liquids containing significantly more errors in children with SLI compared to typically-developing peers and performance on fricatives containing significantly fewer errors in children with SLI compared to their typically-developing peers.

However, the imbalanced distribution of liquids and fricatives within Edwards and Lahey's stimuli prevents any definitive conclusions from their results (Moore, et al., 2010).

Alternatively, Moore et al. (2010) used the Early-, Middle-, and Late-8 consonant groupings with the goal of manipulating the presumed articulatory demands of nonword repetition tasks while other parameters of the stimuli, like syllable structure, duration/length, phoneme predictability and lexicality, were reasonably well controlled. As mentioned previously, they found among a sample of typically-developing college students that performance was significantly worse on the L8NRT compared to performance on the Nonword Repetition Task (NRT; Dollaghan & Campbell, 1998) containing nonwords with only Early-8 and Middle-8 consonants. However, it is possible that the articulatory complexity component was confounded by other factors in the Moore et al. study (e.g., perceptual salience of phonemes) and that, although motoric demands play a role in phoneme acquisition, they have little influence in adults. To support this argument, one might suggest that articulation of all native phonemes becomes so highly skilled and automated that there are no significant differences between early- and late-developing phonemes.

To summarize, this section addressed the natural variation in articulatory demands among English phonemes, and whether this variation contributes to decreased performance in tasks containing early- versus late-developing phonemes. Previous research is inconsistent. On one hand, literature on phonological development suggests that order of phoneme acquisition depends at least in part on the varied motoric demands of phonemes (e.g., Kent, 1992; Stokes & Surendran, 2005); however, there have been inconsistent results when nonword repetition stimuli have varied by such factors (e.g., Edwards & Lahey, 1998). The adult literature is conflicting as well – there were no apparent differences in the DDK rates of syllables containing Early-, Middle-, and Late-8 phonemes (Blumberger, et al., 1995), yet Moore et al. (2010) found decreased nonword repetition performance on stimuli containing only Late-8 phonemes.

2.3.3 Implications for early versus late differences in phonological representation

Whatever causes certain phonemes to be acquired before others (e.g. perceptual factors, motor demands), the fact is that there is a span of several years between mastery of early-learned and late-learned phonemes, a span that could affect the structure of neural networks and the quality and retrieval of phonological representations. Steyvers and Tenenbaum (2005) suggested that early acquired information has the advantage of becoming a "hub" (p. 43) from which other neural connections are established. Later acquired information has decreased centrality, decreased number of connections, and, therefore, decreased utility in networks. If this is the case with phonemes and the organization of phonological representations, then it seems plausible that this effect could continue to affect language processing into adulthood; even with many years of using all sounds, the establishment of the neural networks begins very early, and the structure continues to be reinforced over time. This provides a potential model within typical development of varied quality of phonological representations in long-term memory.

If the early-late phoneme (ELP) effect reflects differences in the quality of the phonological representations, then this could account for differences in nonword repetition. This in turn could recast how the task is viewed, and how poor performance on the task (as occurs in individuals with SLI) is interpreted.

2.4 OBJECTIVES AND RATIONALE OF THE CURRENT PROJECT

The overarching goal of this project was to challenge the assumption that children with SLI have deficits in phonological memory storage capacity simply because they perform poorly on nonword repetition tasks. Many factors are thought to contribute to nonword repetition performance. The most widely addressed, and therefore the focus of this current project, are speech perception demands, articulatory complexity, lexical knowledge, phonological memory storage capacity, and core phonological processing.

In order to determine whether the ELP performance contrast will persist after ruling out other potential explanations, the present work used experimental designs in which the stimulus lists were matched on a variety of potentially important dimensions, and in which different task conditions were employed in a novel “differential diagnosis” approach to tease apart aspects of speech and language processing. This strategy is detailed below.

There are a number of parameters that have been shown to influence reading and other language tasks containing lexical items, such as frequency of occurrence, concreteness, and imageability of the words, to name a few (Juhasz & Rayner, 2006). Furthermore, practically speaking it would be extremely difficult if not impossible to create sufficient lists of real words containing only early-developing or only late-developing phonemes that are also matched on all other parameters. For these reasons, nonwords are used as the focus of testing and analysis for all of the tasks in the study. Although there are still several parameters to control in nonwords, the focus on nonwords maximizes flexibility in stimuli selection and minimizes potential confounds associated with lexical processing.

During auditory presentation of a nonword (as occurs in nonword repetition), phonemes are not all presented at the same time, placing memory storage demands on the participant to remember both the sounds and their serial positions (Baddeley, et al., 1998; Gupta, 2005; Gupta, Lipinski, Abbs, & Lin, 2005). For this project, phonological memory storage capacity demands are considered in two ways. First, all tasks (other than the nonword repetition task in Study 1, Question 1) will use 1-syllable CVC items. Items of this length are well within the memory storage capacity of typical young adults. It could be argued, however, that even 1-syllable items place demands on the memory system when presented aurally, given the serial nature of auditory information. To address this concern, several of the tasks in this project use written stimuli.

It can be difficult to isolate the independent contribution of a targeted factor using direct testing. For example, assessing the contribution of speech perception in nonword repetition by correlating performance on a speech perception and nonword repetition task ignores the fact that other factors (such as memory or lexical knowledge) can be involved in both tasks. In this example it is difficult to determine what shared skill is driving the correlation. An alternative approach is to use a differential diagnosis model of testing – a systematic process of elimination of factors to determine the independent contribution of one factor. Auditory perception and articulatory demands are manipulated in this manner through task selection.

With these ideas in mind, two studies were conducted to investigate the ELP effect in nonword repetition. Study 1 addressed the following specific questions (Table 3):

Question 1: Are the results from the Moore, et al. (2010) study replicable? That is, in normal young adults, are there performance differences in a nonword repetition task between items consisting of early- versus late-developing phonemes?

Question 2: In normal young adults, are there persistent performance differences between items consisting of early- versus late-developing phonemes on language tasks when auditory perceptual demands are eliminated?

Question 3: In normal young adults, are there persistent performance differences between items consisting of early- versus late-developing phonemes on language tasks when overt articulation is eliminated?

Table 3. Study 1 experimental tasks.

Question	**Factor Eliminated**	**Task**	**Input**	**Output**
1	--	Nonword Repetition	Auditory	Spoken
2	Auditory	Nonword Reading	Visual - Written	Spoken
3	Overt Articulation	Lexical Decision	Auditory	Key press

To replicate the results from Moore et al. (2010), addressing the first question in this study, a nonword repetition task was administered to participants using the early- and late-developing phoneme lists selected for this project.

The early and late consonant groups for the current project have been modified from the NRT and L8NRT in two primary ways. First, the late-developing phonemes for the current study only include 7 of the Late-8 consonants. This study includes tasks with visual (i.e., written) stimuli, and because the “soft g” sound (/ʒ/, as in ‘beige’ and ‘measure’) is difficult to represent in orthographic form it was excluded from the set of late developing phonemes. Additionally, the soft ‘g’ does not occur in the initial position of English words and is highly infrequent, thus making it difficult to use in experimental word lists as often as other consonants. Therefore, the late developing consonants in this study will be referred to as the Late-7 (or L7): /s, z, l, r, ʃ, θ, ð/.

The second point of consideration stems from previous studies demonstrating that use of fewer articulatory features within a stimulus list (creating more feature overlap within the list), can affect performance on speech production and working memory tasks (Ellis, 1980; Guediche, Chein, & Fiez, In review; Hintzman, 1967; Levy, 1971; Rogers & Storkel, 1998). Moore et al. (2010) suggested that articulatory feature overlap may have been another potential confound in their study contrasting the NRT and the L8NRT. The consonants from the NRT represent five place features and four manner features of articulation as compared to the L8NRT consonants representing only three place and two manner features. For the current study, the stimuli have been modified to control for feature overlap.

The Late-7 group has the same number of articulatory features as the Late-8 – three place features (interdental, alveolar, palatal) and two manner features (fricative, liquid). The following consonants were selected from Shriberg’s Early-8 and Middle-8 groups to form a subset of 7 early developing phonemes for the experimental tasks in this study: /m, n, p, d, t, f, v/. This Early-7 group (or E7) contains three place features (bilabial, labiodental, alveolar) and three manner features (nasal, stop, fricative) and as such it is now more closely matched in number of features to the L7 group. Voicing is matched as well – each group contains four voiced consonants and three unvoiced consonants.

It is thought that some phonemes are difficult to perceive without the presence of visual cues and/or a verbal context (e.g., Miller & Nicely, 1955). If this is the case, a typical nonword repetition task exacerbates this issue because no visual cues and little to no lexical context (depending on the nonword stimuli design) are available. The challenge in addressing the role of perceptual salience using tasks in which auditory perception is expected to contribute to performance, such as a nonword repetition (Moore, et al., 2010) or phoneme confusability task (Redford & Diehl, 1999), is that we cannot dismiss other potential causes of the performance differences like phonological processing or output demands. Using a process of elimination approach, the second question addressed performance differences between early- and late-developing phonemes when the auditory demands were minimized. In other words, was the perceptual salience of the phonemes driving the ELP effect, or were there differences even when the auditory perceptual factor was minimized? To target this question, a nonword reading task was administered so that the primary input modality was visual (written), placing little to no demand on auditory perception.

Current thinking is that the articulatory demands of phonemes play an important role in phoneme development (Jusczyk, 1992; Kent, 1992; Stokes & Surendran, 2005) as well as in phoneme use in adults (Moore, et al., 2010); however, as mentioned in section 2.3.2.2, due to potential confounds and limited data sets in previous work, the role of articulatory demands merits further investigation. The purpose of Question 3 was to examine the performance differences between early- and late-developing phonemes after the articulatory demands of the task have been minimized or eliminated. In other words, were the articulatory demands of the phonemes driving the early-late phoneme effect, or were there differences even when articulation was reduced? In order to eliminate overt speech production, a lexical decision task was used in which participants were presented with a word or nonword and then made a yes/no decision about whether the stimulus item was a real word or not. For this task, the output response was a yes/no key press; no overt speech production was required.

The aim of Study 2 was to further probe articulatory demands. The following specific questions were addressed (Table 4):

Replication of Study 1, Question 3: In normal young adults, are there persistent performance differences between items consisting of early- versus late-developing phonemes on language tasks when overt articulation is eliminated?

Question 4: In normal young adults, are there persistent performance differences between items consisting of early- versus late-developing phonemes on language tasks when covert articulation is suppressed?

Table 4. Study 2 experimental tasks. A lexical decision task with a 2 x 2 x 2 design, investigating phoneme type (not shown) while manipulating presentation modality and concurrent articulation (CA).

	With/Without Concurrent Articulation	
Presentation Modality (Auditory or Visual)	Auditory without CA	Auditory with CA
	Visual without CA	Visual with CA

Although lexical decision does not involve verbal speech output, previous work has shown that during various working memory and language tasks, participants often carry out an “inner speech” process of planning or rehearsing experimental items in their heads (e.g., Abramson & Goldinger, 1997; Baddeley & Hitch, 1994; D. M. Jones, et al., 2004). A method commonly used to suppress inner speech, or covert articulation, is to instruct participants to use concurrent articulation (CA) while completing the task (Baddeley, 1986). For example, the participant would continually repeat a semantically-neutral, highly-frequent spoken word like "the the the the" during experimental trials. Study 2 was designed to replicate findings in Study 1, Question 3 using a different set of stimuli and new participants. Additionally, Study 2 addressed whether the ELP effect persists after suppressing covert articulation. To examine these two questions, lexical decision tasks were administered with and without CA. If the ELP effect persists in both conditions, this suggests that neither overt articulation nor the processes involved in inner speech production can account for the ELP effect.

As addressed above, the perceptual salience of phonemes and phonological memory storage demands of stimuli presented aurally create potential confounds. Visual (written) presentation of stimuli is not without potential confounds as well, as the written items could require speech-based phonological recoding that is not necessary with auditory presentation (e.g., Baddeley, 2003 as it pertains to verbal working memory). Therefore, both presentation modalities will be used for Study 2. Thus, to determine if performance differences between early- and late-developing phonemes can be observed when articulatory demands are eliminated or suppressed, Study 2 used a 2 x 2 x 2 experimental design – investigating the early-late phoneme contrast in auditory and visual lexical decision, with and without concurrent articulation (as shown in Table 4).

3.0 METHODS

3.1 TECHNICAL SPECIFICATIONS

3.1.1 Stimulus development and presentation.

All auditory stimuli in Study 1 and 2 were digitally recorded in a quiet room by a trained female speaker of Standard American English. An Audio-Technica ATR120 low impedance, dynamic microphone was used at a consistent mouth-to-microphone distance of one inch. Audio files were recorded onto a computer using Adobe Audition 1.5 (44,100 Hz sampling rate and 16-bit resolution).

Auditory stimuli were presented through Sennheiser HD 280 Pro headphones. The presentation volume was set at a comfortable listening level based on the feedback from pilot subjects with normal hearing. During the experimental sessions, the volume was held constant across all participants in Study 1 and 2.

All visual stimuli were displayed in white against a black background. The stimuli were presented in Arial 30-size font, upper case letters in the center of the computer screen. All letters for a given word or nonword were presented simultaneously. Participants were seated approximately 16 inches from the center of the computer monitor and viewed stimuli from the center of their visual fields.

3.1.2 Participant responses.

For verbal responses, participants spoke into an Audio-Technica ATR120 low impedance, dynamic microphone. Responses were recorded digitally onto a computer using Adobe Audition 1.5 (44,100 Hz sampling rate and 16-bit resolution). The microphone was placed on the table directly in front of participants. Because of height and posture differences between participants, the microphone-to-mouth distance ranged from approximately 3 to 9 inches.

During the lexical decision tasks, participants' button press responses were recorded using a serial response (SR) box. The examiner asked all participants to use their two index fingers for button presses. Their left index finger was used for the left-most button on the SR box (labeled 'words') and their right index finger was used for the right-most button (labeled 'nonword'). Their fingers rested on the buttons in between trials.

3.2 STUDY 1 METHODS

3.2.1 Participants.

A recruitment e-mail was sent to a group of undergraduate students at the University of Pittsburgh who had previously participated in a psychology-based reading and language research study. Participants received the recruitment e-mail if they met the following criteria:

1. The participant completed the study in 2007 or later.
2. The participant indicated that s/he was willing to participate in future studies.
3. The participant had valid reading comprehension and vocabulary data available.

4. The participant reported ‘no’ to speaking a non-English native language.
5. The participant reported ‘no’ to speaking a non-English language at home.
6. The participant reported ‘no’ to having a history of reading disorder.
7. The participant reported ‘no’ to having hearing problems.
8. The participant was ≤ 25 years of age at the time of testing.

A total of 30 undergraduate students were enrolled in Study 1 of the current project (4 males). They were 19 or 20 years of age (*M* = 19.37, *SD* = 0.49) with 13 to 15 years of education (*M* = 13.27, *SD* = 0.52 years of education). After consenting to participate, participants reported having no prior or current speech, language, or reading impairment. All participants passed a hearing and vision screening.

3.2.2 Testing procedure.

All procedures were administered individually in a quiet testing room. After consenting to participate, subjects were asked to report any history of speech, language, or reading disorder and then were given brief screenings for vision, hearing threshold, and DDK rates (a measure of articulatory rate capacity).

A hearing threshold screening assessed average pure tone thresholds across 500 Hz, 1 KHz, and 2 KHz. For each frequency level (starting with 1 KHz), a tone was presented at 0 dB and then increased in increments of 10 dB until the participant gave a response indicating that they heard a tone. For reinforcement, the decibel level was increased an additional 10 dB and a tone was presented. At this point, the decibel level was decreased 10 dB until there was no response, then increased 5 dB until the participant responded. The level was recorded. The decibel level was decreased 10 dB again until there was no response, and then increased 5 dB until there was a response. This level was recorded. This process was repeated until there were two responses at the same dB level. Participants passed the hearing screening with average pure tone thresholds across frequencies of 500 Hz, 1 KHz, and 2 KHz that were ≤25 dB.

To demonstrate that participants had adequate visual acuity for experimental tasks, they were asked to read from a near vision test card (i.e. handheld eye chart with letters) that was placed approximately 16 inches from them (J. Schneider, 2002). Participants were able to read with 100% accuracy at least one line (9 letters) in which the letters were smaller than the equivalent of Arial font size 10.

Finally, the three experimental tasks were administered on a computer using the E-prime computer program for psychological experiments (W. Schneider, Eschman, & Zuccolotto, 2002). Nonword repetition was administered first to all participants. Nonword reading and auditory lexical decision followed, with the two tasks counterbalanced across successive participants. Details about the stimuli, administration, and scoring associated with these three tasks are provided below.

3.2.3 Nonword repetition (Study 1, Question 1).

3.2.3.1 Nonword repetition stimuli.

The nonword repetition task in the current study comprised 32 nonwords, 16 nonwords for each phoneme type (E7 and L7), 4 E7 and L7 nonwords at each syllable length (1, 2, 3, and

4). As mentioned previously, the following phonemes comprised the E7 list: /m, n, p, d, t, f, v/. The following phonemes comprised the L7 list: /s, z, l, r, ʃ, θ, ð/. A consonant-vowel (CV) structure was used for all syllables, except the final syllable in which a CVC structure was used. The nonword structures were as follows: CVC, CVCVC, CVCVCVC, CVCVCVCVC.

As described in Dollaghan and Campbell's work (1998), tense vowels are less susceptible to schwa reduction than lax vowels, and are considered to be more easily perceptible. Therefore, the nonword stimuli were constructed using nearly all tense vowels. In order to increase the potential diversity of the nonword stimuli, one lax vowel that is highly perceptible and not prone to schwa reduction, /æ/, was also included in the vowel inventory. Thus, the following vowels were used for this task: /ɑ, æ, e, o, u, ɑʊ, ɑɪ, ɔɪ/.

Dollaghan and Campbell's nonword repetition task (NRT; 1998) has been cited in the literature as a carefully-controlled task (Ellis Weismer, et al., 2000). This current study adopted the strict criteria that were used in creating the NRT and Moore et al.'s L8NRT (2010), and where possible attempted to equate the current E7 and L7 stimuli on additional criteria to minimize potential confounds between lists. Phoneme recurrence, probability of occurrence, and lexicality were the primary factors that were considered (described below and summarized in Table 5).

Phoneme recurrence within a nonword and across the task was considered in order to control phoneme predictability. In the NRT and L8NRT, a phoneme was not used more than one time within a given nonword to reduce phoneme predictability within a given trial. This parameter was also used in the current E7 and L7 lists. Phoneme recurrence across the task was an additional control implemented in the E7 and L7 design that was not factored into the NRT and L8NRT tasks. NRT consonants are used anywhere from 1 to 11 times and L8NRT consonants are used anywhere from 5 to 9 times throughout the task (Moore, et al., 2010). In contrast, all E7 and L7 consonants are used 6 to 10 times in the current nonword repetition task. Table 5 provides a comparison of the NRT-L8NRT and E7-L7 nonword repetition lists on a number of additional parameters relating to phoneme use.

Probability of occurrence and lexicality were considered in the E7 and L7 stimulus development as well. As noted previously, phonotactic probability (the frequency of phoneme occurrence within a language) can also influence performance in tasks like nonword repetition (Edwards, et al., 2004; Munson, Kurtz, et al., 2005; see also section 2.2). Therefore, probability of phonemes and biphones (i.e., segment-to-segment co-occurrence of sounds; Vitevitch & Luce, 2004) was controlled in this task as well. Phonotactic and biphone probability were obtained with a web-based calculator (see Vitevitch & Luce, 2004 for details). Using independent sample *t*-tests to compare the average E7 and L7 probabilities at each syllable length, there were no significant differences between the average phonotactic probabilities and the average biphone probabilities in any position of the two groups of nonwords ($p > 0.08$, see Appendix A, Tables 12 and 13 for values).

To minimize effects of wordlikeness (e.g., Dollaghan, et al., 1993; 1995; see also section 2.2), no syllabic segment corresponded to a Standard American English word. Syllabic segments include all CVs and final CVCs. The constituent CV and VC of the final CVCs were considered as well. There were two exceptions in which a word was used in a syllabic segment – the E7 list contains the CV ‘nah’ and the L7 list contains the VC ‘are’.

Primary stress was assigned to the second syllable of the 4-syllable nonwords and the first syllable of all other nonwords. Stress assignment was validated by a graduate student who was unfamiliar with the task. Average durations (in milliseconds) of the recorded E7 and L7 stimuli were similar at each syllable length ($p > 0.06$, independent sample t-tests, see Appendix A, Table 11 for values).

Table 5. Factors that were considered in the development of the E7 and L7 nonword repetition stimuli lists compared to the NRT and L8NRT.

	NRT vs. L8NRT (Dollaghan & Campbell, 1998; Moore, et al., 2010)		E7 vs. L7 (Current Study)	
Number of items	16 nonwords for each consonant group, 4 nonwords at each syllable length		16 nonwords for each consonant group, 4 nonwords at each syllable length	
Nonword structure	CVC, CVCVC, CVCVCVC, CVCVCVCVC		CVC, CVCVC, CVCVCVC, CVCVCVCVC	
Consonants used*	NRT: 11 unique Early-8 and Middle-8 Cs L8NRT: 8 unique Late-8 Cs		E7: 7 unique Early-8 and Middle-8 Cs L7: 7 unique Late-8 Cs	
Consonant phoneme recurrence across task*	NRT	L8NRT	E7	L7
	/dʒ/ x1	/ʒ/ x5	/p/ x6	/r/ x6
	/k/ x1	/ʃ/ x6	/d/ x7	/l/ x7
	/f/ x2	/z/ x6	/n/ x7	/ð/ x7
	/p/ x3	/s/ x7	/t/ x7	/s/ x8
	/d/ x4	/θ/ x7	/m/ x9	/z/ x8
	/g/ x4	/ð/ x7	/f/ x10	/ʃ/ x10
	/b/ x5	/l/ x9	/v/ x10	/θ/ x10
	/n/ x7	/r/ x9		
	/tʃ/ x9			
	/t/ x9			
	/v/ x11			
Vowels used	Tense vowels, one lax vowel		Tense vowels, one lax vowel	
Vowel phoneme recurrence across task*	NRT	L8NRT	E7	L7
	/ɑ/ x3	/ɑ/ x3	/ɑ/ x6	/ɑ/ x6
	/æ/ x5	/æ/ x5	/æ/ x8	/æ/ x8
	/ɑɪ/ x7	/ɑɪ/ x6	/ɑɪ/ x2	/ɑɪ/ x2
	/ɑʊ/ x5	/ɑʊ/ x9	/ɑʊ/ x6	/ɑʊ/ x6
	/e/ x4	/e/ x4	/e/ x4	/e/ x4
	/i/ x3	/i/ x2	/i/ x0	/i/ x0
	/o/ x4	/o/ x4	/o/ x2	/o/ x2
	/ɔɪ/ x8	/ɔɪ/ x7	/ɔɪ/ x8	/ɔɪ/ x8
	/u/ x1	/u/ x0	/u/ x4	/u/ x4
Phoneme recurrence within item	No repetition of any C or V within a nonword		No repetition of any C or V within a nonword	
CV recurrence across task*	NRT: 18 unique CVs, each used 1-5x L8NRT: 34 unique CVs, each used 1-2x		E7: 20 unique CVs, each used 2x L7: 20 unique CVs, each used 2x	

Table 5 (continued)

CVC recurrence across task*	NRT: 33 unique CVCs, 7 CVCs used 2x L8NRT: 37 unique CVCs, 3 CVCs used 2x	E7: 38 unique CVCs, 2 CVCs used 2x L7: 35 unique CVCs, 5 CVCs used 2x
Cs probability of occurrence	NRT only: Cs only occurred in syllable positions (onset or coda) in which they occurred ≤ 25% of the time in English	---
Phonotactic probability at each syllable length*	---	No significant difference between the E7 and L7 average phonotactic probability in any phoneme position (see Table 12)
Biphone probability at each syllable length*	---	No significant difference between the E7 and L7 average biphone probability in any biphone position (see Table 13)
Lexicality of constituent syllables*	Effort made so no constituent CV or CVC corresponded to a Standard American English word. There were few exceptions in each list. NRT: nigh, chai L8NRT: sow, thy, row, low, ray, lay, shy, sigh, soy, Roy, lows	Effort made so no syllabic segment corresponded to a Standard American English word. (Syllabic segment = all CVs and the final CVCs, including constituent CV and VC of final CVC). There were few exceptions. E7: nah L7: are

* Indicates a factor in which the balance between the early and late phoneme lists has been improved in the current study from previous work with NRT and L8NRT comparisons.

3.2.3.2 Nonword repetition task administration.

Three practice items were administered at the start of the nonword repetition task. All 8 1-syllable nonwords (4 E7 nonwords, 4 L7 nonwords) were administered first, followed by all 2-syllable nonwords, and so on. Nonwords within each syllable length were presented in random order for each participant.

Stimulus presentation was identical to the method described in Moore et al. (2010). At the start of the task, participants read the following instructions on the computer screen: ‘You are going to listen to a series of made-up words presented one at a time. When the fixation point turns green, repeat the made-up word out loud. Say the words exactly the way that you hear them. The first few will be practice. Press any key to begin with the practice items.’ A red fixation cross was displayed on the computer screen 0.5 seconds prior to nonword presentation and remained on the screen throughout the duration of the auditory presentation of a nonword. After the nonword stimulus presentation, the red fixation cross turned green, prompting the participant to provide the spoken response. The green fixation remained on the screen for 2.5 seconds, followed by 0.5 seconds of the red fixation, the next nonword item, and so on. There was a total of 3 seconds between each nonword item.

3.2.3.3 Nonword repetition scoring.

Scoring of the nonword repetition task was completed according to the procedures described by Dollaghan and Campbell (1998). Each phoneme was scored as correct or incorrect in order to compute a percentage of phonemes correct (PPC) at each syllable length and overall. All omission and substitution errors were counted as incorrect; distortions, allophonic variations, and additions were not deducted from the score. Both inter- and intrarater reliability of scoring were measured (reported below). In the few cases (0.2% of all nonword trials) in which a participant stopped the response early and did not repeat a nonword in its entirety, the total number of phonemes repeated correctly was divided by the total number of scoreable phoneme targets.

3.2.4 Nonword reading (Study 1, Question 2).

3.2.4.1 Nonword reading stimuli.

The nonword reading task consisted of 40 nonwords, 20 of each phoneme type (E7, L7; Appendix B, Table 14). All items were 1-syllable, CVC in structure. One primary objective of stimulus development was to ensure that no CVC from this task (nor any constituent CV) was duplicated within the task or in the other tasks in Study 1 (nonword repetition, auditory lexical decision). This parameter was controlled in order to avoid repetition priming effects, i.e. reading latency benefits for repeated items (Coltheart, et al., 2001; Katz, et al., 2005).

In the nonword repetition task, lax vowels were not used because in non-stressed syllables they are susceptible to decreased perceptual salience during listening and vowel reduction during speaking. This is a relevant parameter for the nonword repetition task in which there are multisyllabic items that contain non-stressed syllables. However, for the nonword reading task this parameter is not relevant because 1-syllable items do not contain a non-stressed syllable. To maximize the pool of possible nonwords for the task, both tense and lax vowels were used.

Similar to the nonword repetition task, phoneme recurrence was considered in development of the stimuli. No phoneme was repeated within a nonword. Additionally, Table 15 (Appendix B) shows that phoneme recurrence across the task was generally balanced so that no phonemes are relatively under- or overrepresented.

Many phonological and orthographic factors have been shown to affect reading and lexical decision performance (e.g., Berent, 1997; Berent, Bouissa, & Tuller, 2001; Coltheart, et al., 2001; Perfetti & Hart, 2002; Yates, 2009). For this reason, a number of additional factors were controlled for this task. There were no significant differences between the E7 and L7 nonword lists on the following phonological factors: phonological neighborhood density, weighted phonological neighborhood density based on word frequency, phonotactic probability in each phoneme position, and biphone probability in each position (independent *t*-tests, $t \leq 1.74$, $p \geq 0.09$, see Table 17 in Appendix B for all values). Phonological neighborhood density refers to the number of words that share all but one phoneme with the target nonword (Vaden, Halpen,

& Hickok, 2009). The phonological neighborhood data were obtained using an online calculator from the Irvine Phonotactic Online Dictionary website (see Vaden, et al., 2009). As noted previously, the phonotactic and biphone probabilities were obtained from an online calculator as well (Vitevitch & Luce, 2004).

There were no significant differences between the E7 and L7 nonword lists on the following orthographic factors: number of letters, orthographic neighborhood, mean bigram frequency, and summed bigram frequency by position (independent *t*-tests, $t \leq 2.00$, $p \geq 0.06$; see Table 16 in Appendix B for all values). Orthographic neighborhood refers to the number of words that share all but one letter with the target nonword (while preserving the identity and the positions of the other letters). The bigram frequency refers to the frequency with which two letters in sequence occur within a language (Balota, et al., 2007). For example, one of the items in this nonword reading task is 'fape'. Examples of orthographic neighbors are 'cape' and 'fade'. 'Fape' has three bigrams – fa, ap, pe. The mean bigram frequency for this item is the sum of each of the three bigram frequencies divided by the total number of bigrams (three in this example). Bigram frequency by position considers the frequency of each bigram only within its position in the target (non)word. In the nonword 'fape', for example, the bigram frequency by position only considers the frequency that 'fa' occurs as the first bigram of a word. These orthographic measures were computed using the English Lexicon Project website (Balota, et al., 2007).

A final consideration was the constituent components of the nonwords. In order to assess the wordlikeness of the CVs, judgment ratings were obtained from nine native English speakers who were blinded to the purpose of the study. After the CV unit of each nonword was read aloud to the raters, they were asked to give a rating from 1 to 5, 1 being not at all wordlike and 5 being a real word. There were no significant differences in average wordlikeness ratings between the

CVs in the E7 and L7 nonword lists (M_{E7} = 3.54 (*SD* = 1.47), M_{L7} = 2.76 (*SD* = 1.60), independent sample *t*-test, $t(38) = 1.61$, $p = 0.12$).

Previous work has shown that the rhyme, i.e. the VC units in the current word lists, has the greatest influence on pronunciation (Jared, McRae, & Seidenberg, 1990; Treiman, Goswami, & Bruck, 1990; Treiman & Zukowski, 1988). The consistency of the nonwords, or the degree to which a (non)word has the same pronunciation as similarly spelled words (Plaut, McClelland, Seidenberg, & Patterson, 1996), was considered using the rhymes of the nonwords in the current nonword reading task. In a nonword reading task, Treiman and colleagues (1990) found performance differences when the spelling pattern for the VC unit was shared with many words versus few or no words. For this reason, the number of friends (words with a shared spelling pattern and pronunciation) was compared between the E7 and L7 items in this task. No significant differences were found (M_{E7} = 6.60 (*SD* = 5.60), M_{L7} = 6.35 (*SD* = 6.10), independent sample *t*-test, $t(38) = 0.14$, $p = 0.89$).

A consistency ratio was also computed in order to account for the effects of both friends and enemies (words that have the same spelling pattern but different pronunciation; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Pattamadilok, Knierim, Kawabata Duncan, & Devlin, 2010; Plaut, et al., 1996). The number of friends was divided by the sum of the total number of friends and enemies (Pattamadilok, et al., 2010). There were no significant differences in the consistency ratio between the E7 and L7 lists (M_{E7} = 0.99 (*SD* = 0.06), M_{L7} = 0.96 (*SD* = 0.08), independent sample *t*-test, $t(31) = 1.05$, $p = 0.30$). Note that the number of friends and the consistency ratio were considered separately because of the unique information that each measure provides. The consistency ratio factors in the effect of enemies, but for consistent words with no enemies the ratio is 1. This is true if a target nonword has 2 friends or 20 friends. The use of a separate measure to compare the number of friends was motivated by the findings of Treiman and colleagues (1990).

3.2.4.2 Nonword reading task administration.

Three practice items were administered at the start of the nonword reading task. All 40 nonwords were presented in random order for each participant. Prior to the practice trials, participants read the following instructions on the computer screen: ‘You will see a series of made-up words presented on the screen one word at a time. Read each made-up word as quickly and accurately as possible. When the '+' appears, press the spacebar to get the next item. You will start with some practice items. Press the spacebar when you are ready to begin.’ A white fixation cross appeared and remained on the computer screen until the participant hit the space bar to elicit a nonword. When the space bar was pressed, the screen went blank for 250 ms, and then the nonword appeared and remained on the screen until the participant responded aloud. The nonword item turned light gray to indicate that the microphone registered the voice response (if it did not register the response, the examiner pressed a button to advance the nonword from white to light gray). Via button press, the examiner then recorded two pieces of information: first, whether the voice response time was registered correctly (score of 1) or not (0), and second, whether the participants’ pronunciation was correct (1) or not (0). With the examiner’s last button press, a white fixation cross appeared until the participant pressed the space bar to cue the next nonword, and so on. After the initial instructions and then again after the practice items, the participants were given an opportunity to ask questions before proceeding with the task.

3.2.4.3 Nonword reading scoring.

Two measurements were recorded for this task: whole-word accuracy and reading latency. Using a strict scoring system for accuracy, responses were marked as correct if the participant pronounced the nonword identically to the target pronunciation (see Appendix B, Table 14 for pronunciations). In a lax scoring system for accuracy, two accommodations were made from the strict scoring. First, a pronunciation was scored as correct if the participant used any legal pronunciation for a given rhyme. Second, because of the ambiguity in determining the need for voiced or voiceless 'th' in the initial position of nonwords, in the lax scoring either phoneme was scored as correct when used in the initial position of any 'th' nonword. Both inter- and intrarater reliability of scoring were measured (reported below).

For example, the target pronunciation for the nonword 'thear' was /θir/ (voiceless 'th', rhymes with 'dear'). Using the strict scoring, /θir/ was the only pronunciation that received a correct score. Using the lax scoring, the nonword could be read as /θir/ or as /θɛr/ (rhymes with 'bear') since those are both legal pronunciations for the –ear word body. Pronunciations using the voiced 'th' (e.g., /ðir/) were also scored as correct.

Reading latency was measured as the duration from the appearance of the nonword to the start of phonation of the response. Response time was recorded using a microphone attached to an E-prime SR box; however, during analysis it was noted that response time sensitivity was not consistent across speakers or phonemes (e.g., response time was not marked until vowel phonation for some nonwords with initial-position fricatives).

To obtain a more reliable measure of reading latency, response times were measured using the spectral and waveform views in Adobe Audition. Appearance of the nonword was marked by the offset of a high-pitched beep that was played while the 250-millisecond blank screen was displayed during the task (the speaker volume was muted so that the beep could not be heard by a listener). During scoring, the beeps were isolated using a band pass filter in Adobe Audition, and their onset and offset times were marked using the “Find Phrases and Mark” function. The same function was applied to the digital recordings of participants’ responses. Using the spectral view, the phrase mark was adjusted so that it was set at the start of phonation. This was verified by using the phrase playback feature, which played the audio for only the selected phrase. For example, if the phrase mark was placed slightly after the start of phonation for a response, during playback the nonword sounded ‘clipped,’ indicating that an adjustment needed to be made. Reliability of this procedure is reported below. The reading latencies obtained manually were used in the analysis.

3.2.5 Auditory lexical decision (Study 1, Question 3).

3.2.5.1 Auditory lexical decision stimuli.

The auditory lexical decision task comprised 30 nonwords and 30 words, all 1-syllable, CVC in structure. For the nonwords, there were 15 of each phoneme type (E7, L7). For the words, there were 15 E7, 12 L7, and 3 ‘mixed’ stimuli. ‘Mixed’ words contained an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited number of real CVC words that consist of only L7 phonemes and that meet all of the other criteria listed below. These ‘mixed’ words were included in the calculations of phoneme recurrence, but were not included in any analyses with E7 and L7 comparisons.

As described above for the nonword reading task, both tense and lax vowels were used in the auditory lexical decision task to maximize the pool of possible nonwords. Additionally, to avoid repetition priming effects, an effort was made to use CVCs (and constituent CVs) that were not used in the nonword repetition or nonword reading tasks. One exception is the nonword /θir/, which occurred in the nonword reading and the auditory lexical decision task.

Table 18 (Appendix C) shows the nonword stimuli for the auditory lexical decision task. Many parameters that were controlled in the nonword reading stimuli were also controlled here (phoneme recurrence: see Appendix C, Table 19; phonological factors: independent *t*-tests, $t \leq 1.69$, $p \geq 0.10$, see Appendix C, Table 20 for all values). It has been suggested that orthographic information can influence phonological representations as well (e.g., Taft, 2006). However, because the primary consideration for the auditory stimuli was to control for repetition effects, phoneme recurrence, and phonological factors, the auditory lexical decision stimuli were not controlled for orthographic factors. This limitation is addressed in Study 2, in which all auditory and visual lexical decision stimulus lists are controlled for both phonological and orthographic factors.

In contrast to the nonword repetition task in which constituent syllables were primarily nonwords, the constituent CVs of the auditory lexical decision task were designed to be more wordlike to increase the difficulty of determining whether each item was a word or nonword. The focus was to control the CVs so that the participant would be more likely to listen to all three phonemes before making a decision. Twelve of the 15 E7 nonwords contained CV words, and 11 of the 15 L7 nonwords contained words. In each group, there was word frequency data available for 11 of the CVs (from the HAL word frequency database on the English Lexicon Project website). There was no significant difference in word frequency between the CVs in E7 and L7 nonword lists (log frequency: M_{E7} = 8.90 (SD = 1.99), M_{L7} = 9.56 (SD = 2.29), independent sample t-test, $t(20) = 0.72$, $p = 0.48$). Average durations (in milliseconds) of the recorded E7 and L7 stimuli were not significantly different (M_{E7} = 709.40 (SD = 74.12), M_{L7} = 738.27 (SD = 50.03), independent sample t-test, $t(28) = 1.25$, $p = 0.22$).

Although the nonword stimuli are the focus of analysis, an effort was made to control several factors within the word lists as well, so that lexical decisions for nonwords could be made within a similar word-nonword environment. Using the nonwords as the basis for stimulus construction, a real word was created by changing either the vowel or the final consonant of a nonword (note one exception, the word 'sill', in which both the vowel and the final consonant were changed). Phoneme recurrence was considered (Appendix C, Table 19), and there were no significant differences between the E7 and L7 word lists on the phonological variables mentioned previously (independent t-tests, $t \leq 0.99$, $p \geq 0.33$, see Appendix C, Table 20 for all values). Likewise, there was no significant difference in word frequency between the E7 and L7 word lists (log frequency: M_{E7} = 8.68 (SD = 2.17), M_{L7} = 8.31 (SD = 2.61), independent sample t-test, $t(25) = 0.40$, $p = 0.69$). Average durations (in milliseconds) of the recorded E7 and L7 stimuli were similar (M_{E7} = 709.87 (SD = 88.89), M_{L7} = 739.25 (SD = 69.97), independent sample t-test, $t(25) = 0.94$, $p = 0.36$).

3.2.5.2 Auditory lexical decision task administration.

Four practice items (two words, two nonwords) were presented in random order at the start of the task. Following the practice, all 60 experimental items were presented in random order for each participant.

At the start of the task, participants read the following instructions on the computer screen: ‘You will hear a series of words and made-up words presented one item at a time. If the item is a word, press the button on the left marked 'word'. If the item is a made-up word, press the button on the right marked 'nonword'. Respond as quickly and accurately as possible. You will start with some practice items. Press any button when you are ready to begin.’ A white fixation cross appeared and remained on the screen throughout the duration of the task (in order to provide the participants with a focal point that was similar to the nonword reading task). After 500 ms, the first trial was presented; the participant had unlimited time to respond. When a response was provided, there was a 1500 ms intertrial interval, the next trial presentation with unlimited time to respond, a 1500 ms intertrial interval, and so on. Response parameters are described in section 3.1.2.

3.2.5.3 Auditory lexical decision scoring.

Two measurements, accuracy and reaction time (RT), were recorded. Accuracy was recorded as correct (score of 1) if the participant correctly indicated an item was a word or nonword or incorrect (score of 0) if they did not. Based on a typical scoring procedure used in the literature (e.g., Goldinger, 1996; Pattamadilok, et al., 2010; Ziegler & Ferrand, 1998), RT was recorded as the duration from the onset of the stimulus item to the onset of a participant response. However, Goldinger (1996) discusses the concern that, when using the stimulus onset to calculate RT, results are confounded if the mean durations of the items within the two stimuli lists differ. This concern will be addressed further in a secondary analysis.

3.2.6 Scoring reliability.

Interrater and intrarater reliability measures were obtained for tasks requiring judgments of accuracy – nonword repetition and nonword reading in Study1. An undergraduate research assistant independently scored tasks for interrater reliability using participants' digital audio files. A subset of six participants (20% of the sample for each subset) was randomly selected for each task. For the E7 items in the nonword repetition task, phoneme-by-phoneme percentages of agreement for judgment of correctness ranged from 96% to 99%, with an average of 98%. Phoneme-by-phoneme percentages of agreement for judgment of correctness for the L7 items ranged from 91% to 94%, with an average of 93%. For the E7 items in the nonword reading task, percentages of agreement for judgment of correct CVC pronunciation ranged from 95% to 100%, with an average of 98%. Percentages of agreement for the L7 items ranged from 90% to 100% with an average of 94%.

To measure intrarater reliability for the nonword repetition and nonword reading tasks, the primary scorer (the author) randomly selected two different subsets of six participants' digital audio files to re-score (one subset per task). The second round of scoring was completed approximately five months after the initial scoring and the author was blinded to the participants' original scores. For the E7 items in the nonword repetition task, phoneme-by-phoneme percentages of agreement for judgment of correctness ranged from 98% to 100%, with an average of 99%. Phoneme-by-phoneme percentages of agreement for judgment of correctness for the L7 items ranged from 92% to 99%, with an average of 97%. For the E7 items in the nonword reading task, percentages of agreement for judgment of correct CVC pronunciation was 100%. Percentages of agreement for the L7 items ranged from 90-100%, with an average of 97%.

To measure reliability of the procedure used to obtain reading latencies for the nonword reading task, an undergraduate research assistant independently marked the onset of phonation for the responses of four randomly selected participants. The average reading latencies of these four participants were nearly identical between the two scorers (M_1 = 681.0 ms (*SD* = 78.3), M_2 = 679.3 ms (*SD* = 77.9)).

3.3 STUDY 2 METHOD

3.3.1 Participants.

A second group of 20 undergraduate students was enrolled in Study 2 (8 males). Participants were recruited using the same method and criteria as for Study 1 (see section 3.2.1). They were 18 to 22 years of age (*M* = 18.75, *SD* = 1.12) with 12 to 15 years of education (*M* = 12.50, *SD* = 0.83 years of education). Participants were native English speakers who reported having no prior or current speech, language, or reading impairment. One participant reported receiving short-term, phonics-based reading assistance in elementary school; because he did not receive a diagnosis for reading impairment and did not have extended school intervention, he was included in the study. During data analysis, his response patterns were compared to other participants and it was noted that his performance conformed to that of the rest of the group. All participants demonstrated adequate visual acuity for experimental tasks and passed the hearing screening with average pure tone thresholds that were ≤25 dB.

3.3.2 Testing procedure.

All procedures were administered individually in a quiet testing room. After consenting to participate, subjects were asked to report any history of speech, language, or reading disorder and then were given brief screenings for vision, hearing threshold, and DDK rates. Finally, the experimental tasks were administered on the computer using E-prime (W. Schneider, et al., 2002). Each participant completed four experimental conditions for the lexical decision task: auditory lexical decision, auditory lexical decision with CA, visual lexical decision, and visual lexical decision with CA.

The task conditions were administered in pseudorandom order such that each subject was randomly assigned to a fixed condition order. The orders were designed so that each task condition was administered as the first condition five times, as the second condition five times, and so on. There was one exception due to examiner error: The order for auditory lexical decision with CA and the visual lexical decision conditions was switched for one participant. As a result, auditory lexical decision is the first task condition six times and the second condition four times across participants; visual lexical decision is the first task condition four times and the second condition six times. All other order positions for all tasks are counterbalanced.

3.3.3 Stimuli.

Four word-nonword lists were used in the lexical decision tasks employed in Study 2 (see Tables 21-24 in Appendix D for stimuli lists). All items are 1-syllable, CVC in structure. No word or nonword was repeated across the lists.

Each nonword list consisted of 30 nonwords, 15 of each phoneme type (E7, L7). Phoneme recurrence was considered across lists (see Appendix D, Table 25). The phonological and orthographic factors addressed in Study 1 (including the consistency ratio and number of friends) were controlled in Study 2 as well. There were no significant differences by phoneme type across the four lists for any of the controlled factors (phonological factors: $F \leq 3.88$, $p \geq 0.05$; orthographic factors: $F \leq 3.62$, $p \geq 0.06$, see Tables 27 and 28 in Appendix D for all values). Additionally, average durations (in milliseconds) of the recorded E7 and L7 stimuli were similar across the four lists ($F \leq 3.11$, $p \geq 0.08$, see Table 31 in Appendix D for all values).

As mentioned previously, an effort was made to control several factors across the word stimuli so that lexical decisions for nonwords could be made within a similar word-nonword environment. Each of the four lists consists of 30 CVC words: 15 E7 words, 12 L7 words, and 3 'mixed' words. See Table 26 (Appendix D) for phoneme recurrence within the lists.

Most of the phonological and orthographic factors were controlled by phoneme type across the four lists as well (phonological factors: $F \leq 3.80$, $p \geq 0.05$; orthographic factors: $F \leq 3.09$, $p \geq 0.08$, see Tables 29 and 30 in Appendix D for all values). Two exceptions were weighted phonological neighborhood (main effect of phoneme type, $F = 5.23$, $p = 0.02$) and phonotactic probability of the third phoneme (main effect of phoneme type, $F = 4.58$, $p = 0.04$). Word frequency as indexed by both the Kucera-Frances and HAL databases was an additional factor that was controlled for the word stimuli, with no significant differences found between the comparisons ($F \leq 1.30$, $p \geq 0.28$, see Table 32 in Appendix D for all values). Finally, average durations (in milliseconds) of the recorded E7 and L7 stimuli were similar across the four lists ($F \leq 1.53$, $p \geq 0.21$, see Table 31 in Appendix D for all values).

3.3.4 Task administration.

Stimuli lists were counterbalanced across tasks. The four lists were assigned to the four task conditions in pseudorandom order such that each list was randomly assigned to a fixed task condition order. The list assignment was designed so that every subject received each list one time, and across subjects each list occurred in every task five times.

Prior to the start of the tasks, participants were given an opportunity to practice concurrent articulation. “One, two, three, four” was the phrase selected to articulate because it has a relatively well-balanced use of both early-developing and late-developing phonemes and has been shown to produce similar effects as other commonly used CA phrases (Baddeley, 1986). Participants rehearsed articulating the “one, two, three, four” phrase approximately one time per second. First they practiced with an online metronome, and then they practiced without the metronome to demonstrate that they could maintain the approximate pace.

All four task conditions were administered similarly to the procedures described for the auditory lexical decision task in Study 1 (Question 3). Instructions were modified accordingly for the tasks using visual stimuli instead of auditory stimuli and for the tasks with CA. One change from the procedures in Study 1 was the insertion of a small break halfway through each task. The break was added so that participants would have shorter intervals of CA, though the break was included in all tasks (even those without CA) to maintain uniformity across the tasks.

3.3.5 Scoring.

Two measurements, accuracy and RT, were recorded with the SR box for the lexical decision tasks. Accuracy was recorded as correct (score of 1) if the participant correctly indicated an item was a word or nonword or incorrect (score of 0) if they did not. RT was recorded as the duration from the onset of the stimulus item to the onset of a participant response. In the few cases (0.2% of all trials) in which a participant was not ready for a task to begin or for a task to resume after a break, the RT for that trial was not included in the analysis.

4.0 RESULTS

Several accounts have been developed to address the phonological processing deficits in communication disorders like SLI. However, confounding factors make it difficult to determine the independent contribution of various phonological skills in the tasks used to inform these theoretical accounts. The specific goal of this project was to test the hypothesis that differences in nonword repetition performance cannot be taken as evidence of a specific problem in phonological storage capacity using a manipulation of early- and late-developing phonemes. The research questions focused on: 1) replicating previous findings of early-late phoneme differences in nonword repetition (Study 1 via a nonword repetition task), 2) examining early-late phoneme differences after minimizing auditory perceptual factors (Study 1 via a nonword reading task, and Study 2 via auditory and visual lexical decision tasks), 3) examining early-late phoneme differences after minimizing articulatory demands (Study 1 and Study 2 via lexical decision tasks performed in silence), 4) and examining early-late phoneme differences after minimizing covert articulation (Study 2 via a lexical decision task performed with CA).

4.1 NONWORD REPETITION (STUDY 1, QUESTION 1)

The first research question in Study 1 addressed whether the early-late phoneme contrast from Moore et al. (2010) was replicable. That is, are there performance differences in nonword repetition for items containing early-developing versus late-developing phonemes? The nonword repetition task in this work is positioned to extend the prior work as well. As mentioned previously (section 3.2.3), the E7 and L7 nonword stimuli were designed to control for some factors that were not considered when developing the NRT (Dollaghan & Campbell, 1998; Moore et al., 2010) and L8NRT (Moore et al., 2010), such as articulatory feature overlap and phonotactic probability.

A 2 x 4 (phoneme type x syllable length) within-subjects ANOVA was used to analyze participants' performance on the nonword repetition task. Main effects were significant for each factor, indicating that performance was significantly decreased for L7 stimuli compared to E7 stimuli ($F(1,29) = 72.41$, $p < 0.001$, partial $\eta^2 = 0.71$), and significantly different across syllable lengths ($F(3,87) = 37.49$, $p < 0.001$, partial $\eta^2 = 0.56$). These findings for syllable length were analyzed further with post-hoc pairwise comparisons using paired sample *t*-tests, corrected for multiple comparisons. These results indicated that, collapsed across phoneme type, there was no significant difference in performance between 1-syllable and 2-syllable items ($p = 0.25$). Performance significantly differed for all other syllable length contrasts, however, with performance decreasing as the number of syllables increased ($p \leq 0.004$).

There was a significant phoneme type x syllable length interaction ($F(3,87) = 10.14$, $p < 0.001$, partial $\eta^2 = 0.26$) indicating that performance on L7 items was significantly decreased compared to E7 items as a function of syllable length. As predicted, participants scored significantly lower on L7 nonwords at all syllable lengths and overall compared to E7 nonwords (one-tailed paired sample *t*-tests, $t \geq 3.30$, $p \leq 0.003$, corrected for multiple comparisons; see Table 6 for all values). This size of the phoneme type difference grew from just under 4% for 1-syllable items to just over 14% for 4-syllable stimuli. The magnitude of the performance difference was large for all comparisons (Cohen's $d \geq 0.86$). These results are mostly consistent with the Moore et al. findings in which there was a large effect size at all syllable lengths and overall (Cohen's $d \geq 1.18$), with the exception of the one-syllable nonword scores (Cohen's d = 0.01).

Table 6. Mean (*M*), standard deviation (*SD*), paired *t*-test (*t*), *p*-value (*p*), and Cohen's *d* (*d*) of scores for the nonword repetition task at each syllable length.

Score	E7 *M* (*SD*)	L7 *M* (*SD*)	Test Statistics
1PPC	98.06 (3.58)	94.17 (5.43)	$t(29) = 3.30$ $p = 0.002^*$ $d = 0.86$
2PPC	98.50 (2.68)	92.00 (5.81)	$t(29) = 6.36$ $p < 0.001^*$ $d = 1.46$
3PPC	95.71 (6.79)	86.41 (9.66)	$t(29) = 4.44$ $p < 0.001^*$ $d = 1.13$
4PPC	90.19 (7.81)	75.91 (11.71)	$t(29) = 8.23$ $p < 0.001^*$ $d = 1.46$
TOTPPC	94.24 (3.89)	84.64 (6.89)	$t(29) = 9.21$ $p < 0.001^*$ $d = 1.75$

Note. Scores are reported as percentage of phonemes correct (PPC) for stimuli 1, 2, 3, and 4 syllables in length and for the task overall (TOTPPC). E7 = Early-7 stimuli; L7 = Late-7 stimuli.
* Indicates significance at $p \leq 0.01$, correcting for multiple comparisons

4.2 NONWORD READING (STUDY 1, QUESTION 2)

The second experimental question was whether an ELP effect can be observed when auditory perceptual factors are minimized or eliminated. To target this question, a nonword reading task was administered in which there was no auditory component to the task. Nonword items were presented visually, and participants were asked to read the items aloud as quickly and accurately as possible.

Both a strict scoring and a lax scoring procedure were implemented in this study because of the ambiguity in pronouncing a few of the nonwords. The two scoring procedures produced a similar pattern of results for whole-word accuracy and RTs, therefore only data from the lax scoring procedure will be presented here (see Appendix E, Table 33 for the results based on the strict scoring).

Based on two-tailed paired *t*-tests to compare participants' whole-item accuracy and RT for E7 and L7 nonword items, the results show that participants were significantly less accurate reading nonword items containing late-developing phonemes (p = 0.004, see Table 7). There were no significant differences in participants' RTs between the early and late phoneme groups.

Table 7. Mean (*M*), standard deviation (*SD*), paired *t*-test (*t*), *p*-value (*p*), and Cohen's *d* (*d*) of whole-item accuracy and RT for the nonword reading task (Study 1).

Dependent Measure	E7 *M* (*SD*)	L7 *M* (*SD*)	Test Statistics
Nonwords (Accuracy)	79.00 (13.98)	72.67 (14.55)	$t(29) = 3.14$ $p = 0.004$ $d = 44.38$
Nonwords (RT in ms)	734.45 (195.26)	736.51 (172.10)	$t(29) = 0.16$ $p = 0.88$ $d = 0.01$

Note. RT is reported for correct responses only. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

4.3 AUDITORY LEXICAL DECISION (STUDY 1, QUESTION 3)

The third experimental question of Study 1 was whether an ELP effect can be observed when articulatory demands are minimized or eliminated. To target this question, an auditory lexical decision task was used in which no spoken response was required. On average across all trials, participants were 86% accurate in their responses (range 72% to 97%).

As mentioned previously, performance on nonword items is the primary focus in this project. Based on two-tailed paired sample *t*-tests to compare participants' accuracy and RT for E7 and L7 nonword items, the results show that participants were both significantly less accurate and significantly slower on nonword items containing late-developing phonemes ($p \leq 0.001$, see Table 8). The effect size for accuracy was large ($d = 120.60$), but small for RT ($d = 0.35$; Cohen, 1977). Although the word items were not the primary focus of the analysis, the same general pattern was observed (accuracy $d = 107.90$, RT $d = 0.60$; see Appendix E, Table 34).

In auditory lexical decision, Goldinger (1996) discusses the concern about the common practice of measuring RT from stimuli onset to onset of output when the duration of stimuli varies between the experimental lists. As indicated above (section 3.2.5), the durations of the E7 and L7 nonwords were not significantly different ($M_{E7} = 709.87$ ($SD = 88.89$), $M_{L7} = 739.25$ ($SD = 69.97$)). However, the difference in actual value of the means (29.38 ms) constitutes nearly half of the difference in RT between E7 and L7 nonword stimuli (63.87 ms). For this reason, a secondary analysis was performed to adjust for duration of stimuli. The length of the recorded stimuli was subtracted from the original RT for each trial in order to establish an adjusted RT. Using a paired sample *t*-test, the average adjusted RT was not significantly different between E7 and L7 nonwords (p = 0.20, see Table 8). In contrast, the average adjusted RT between E7 and L7 words was significantly different (p = 0.003, see Table 34 in Appendix E).

To examine the discriminability between word and nonword stimuli, the signal detection measure d' was computed for E7 and L7 items according to the procedures described in Stanislaw and Todorov (1999). Correct responses (or discriminations) to word items were scored as hits, and incorrect discriminations of nonword items were scored as false alarms. The discriminability measure was defined as: d' = Z(hits) – Z(false alarms), where Z is the standardized value of hits and false alarms based on a normal distribution. Based on a paired sample t-test, the discriminability of L7 items compared to E7 items was significantly lower ($p < 0.001$, see Table 8). To measure response bias for words and nonwords, β was also calculated according to the procedures described in Stanislaw and Todorov. Based on a paired sample t-test, there was no significant difference in β between the E7 and L7 lists (M_{E7} = 1.83 (SD = 1.73), M_{L7} = 1.17 (SD = 0.64), $t(29) = 2.04$, p = 0.05).

Table 8. Mean (M), standard deviation (SD), paired t-test (t), p-value (p), and Cohen's d (d) of accuracy, RT, Adjusted RT, and signal detection measure (d') for the auditory lexical decision task (Study 1).

Dependent Measure	E7 M (SD)	L7 M (SD)	Test Statistics
Nonwords (Accuracy)	91.56 (7.20)	82.00 (8.60)	$t(29) = 6.02$ $p < 0.001$ $d = 120.60$
Nonwords (RT in ms)	1161.74 (172.96)	1225.61 (194.28)	$t(29) = 3.87$ $p = 0.001$ $d = 0.35$
Nonwords (Adjusted RT in ms)	458.76 (175.15)	480.45 (193.75)	$t(29) = 1.30$ $p = 0.20$
d'	2.86 (0.56)	1.85 (0.70)	$t(29) = 8.37$ $p < 0.001$

Note. RT is reported for correct responses only. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

4.4 LEXICAL DECISION: MODALITY AND ARTICULATION INFLUENCES (STUDY 2)

The purpose of Study 2 was to replicate the findings in the lexical decision task in Study 1 and to extend these results by considering a fourth experimental question, whether an ELP effect can be observed when covert articulation is minimized. This study is also designed to provide another manipulation of auditory perceptual demands, which was previously addressed using the nonword reading task (Study 1, Question 2).

A 2 x 2 x 2 (phoneme type x presentation modality x CA) within-subjects ANOVA was used to analyze participants' performance on accuracy and RT. On average across all trials, participants were 91% accurate in their responses (range 85% to 94%). The results for nonword performance are listed in Table 9. For nonword accuracy, main effects were significant for each factor ($F(1,19) \geq 4.83$, $p \leq 0.04$) indicating that performance was significantly better for E7 stimuli compared to L7 stimuli, for visual items compared to auditory items, and without CA compared to performance with CA. There were no significant interactions ($p \geq 0.25$). For average nonword RTs, a significant main effect was found for presentation modality ($F(1,19) = 356.63$, $p < 0.001$), but not for phoneme type or concurrent articulation ($F(1,19) \leq 0.59$, $p \geq 0.45$). There were no significant interactions for nonword RT ($p \geq 0.21$). An adjusted RT was calculated using a procedure identical to the one described in section 4.3. The pattern of results for the adjusted RT was the same as the pattern observed for RT (see Table 9).

To examine the discriminability and response bias between word and nonword stimuli, the signal detection measures *d'* and β were computed. These measures were calculated using the same procedures described in section 4.3. Separate 2 x 2 x 2 within-subjects ANOVAs were used to analyze *d'* and β. The pattern of results for *d'* is consistent with the results for participants' nonword accuracy. That is, there were significant main effects for each factor ($F(1,19) \geq 8.57$, $p \leq 0.01$), but there were no significant interactions ($p \geq 0.16$, see Table 9). For response bias, there was a significant main effect of modality ($M_{Auditory} = 2.33$ ($SD = 2.01$), $M_{Visual} = 3.35$ ($SD = 2.40$), $F(1,19) = 12.12$, $p = 0.003$, partial $\eta^2 = 0.39$), but there were no significant main effects of phoneme type or CA ($F(1,19) \leq 2.57$, $p \geq 0.13$). There were no significant interactions of response bias ($p \geq 0.08$).

Table 9. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), *p*-value (*p*), and partial eta squared (η^2) of accuracy, RT, Adjusted RT, and signal detection measure (*d'*) for the nonword items in lexical decision (Study 2).

Statistical Effects	Nonword Accuracy	Nonword RT (ms)	Nonword Adjusted RT (ms)	*d'*
M_{E7} (*SD*)	96.08 (6.85)	901.19 (259.93)	528.73 (186.62)	3.05 (0.70)
M_{L7} (*SD*)	92.33 (9.37)	909.23 (276.78)	529.20 (191.23)	2.80 (0.72)
$ME_{PhonemeType}$	$F(1,19) = 8.02$ $p = 0.01$ partial $\eta^2 = 0.30$	$F(1,19) = 0.59$ $p = 0.45$ partial $\eta^2 = 0.03$	$F(1,19) = 0.002$ $p = 0.97$ partial $\eta^2 < 0.001$	$F(1,19) = 8.57$ $p = 0.01$ partial $\eta^2 = 0.31$
M_{Aud} (*SD*)	90.67 (9.10)	1140.20 (124.90)	389.11 (125.88)	2.54 (0.69)
M_{Vis} (*SD*)	97.75 (5.80)	670.22 (128.66)	668.83 (126.23)	3.28 (0.54)
$ME_{PresentationModality}$	$F(1,19) = 41.92$ $p < 0.001$ partial $\eta^2 = 0.69$	$F(1,19) = 356.63$ $p < 0.001$ partial $\eta^2 = 0.95$	$F(1,19) = 127.81$ $p < 0.001$ partial $\eta^2 = 0.87$	$F(1,19) = 71.77$ $p < 0.001$ partial $\eta^2 = 0.79$
M_{noCA} (*SD*)	95.50 (6.94)	905.85 (278.01)	530.58 (184.44)	3.08 (0.69)
M_{CA} (*SD*)	92.92 (9.51)	904.57 (258.26)	527.35 (193.32)	2.74 (0.72)
ME_{CA}	$F(1,19) = 4.83$ $p = 0.04$ partial $\eta^2 = 0.20$	$F(1,19) = 0.01$ $p = 0.94$ partial $\eta^2 < 0.01$	$F(1,19) = 0.04$ $p = 0.84$ partial $\eta^2 = 0.002$	$F(1,19) = 13.37$ $p = 0.002$ partial $\eta^2 = 0.41$
Interactions	No Interactions $p \geq 0.25$	No Interactions $p \geq 0.21$	No Interactions $p \geq 0.21$	No Interactions $p \geq 0.16$

Note. RT is reported for correct responses only. A 2 x 2 x 2 within-subjects ANOVA design comparing phoneme type x presentation modality x presence of concurrent articulation was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; Aud = Auditory presentation; Vis = Visual presentation; noCA = without concurrent articulation; CA = with concurrent articulation.

There was no significant effect of phoneme type for real word accuracy, a pattern of results which differed from that for the nonword items ($p = 0.48$; see Table 35 in Appendix E for results for the word items).

Table 10 lists the average accuracy on E7 and L7 items for each lexical decision task. To further examine the performance by phoneme type for each task, the differences in average accuracy between the E7 and L7 items were computed. Importantly, there were differences in performance between E7 and L7 items for each task condition. The difference was smallest when neither auditory stimuli nor CA were used (2% difference between E7 and L7 items on the visual lexical decision task without CA). The lack of a significant interaction indicates that the smaller difference in the visual lexical decision without CA condition is not reliable; however, it is possible that a ceiling effect in the visual modality contributed to the lack of significant interaction (as elaborated in the general discussion).

Table 10. Mean (*M*), standard deviation (SD), and E7-L7 difference in accuracy for the nonword items in each lexical decision task for Study 2.

Task	E7 *M* (*SD*)	L7 *M* (*SD*)	E7 – L7
Auditory without CA	94.67 (8.81)	90.67 (7.62)	4.00%
Auditory with CA	91.00 (7.26)	86.33 (10.92)	4.67%
Visual without CA	99.33 (2.05)	97.33 (3.99)	2.00%
Visual with CA	99.33 (2.98)	95.00 (9.88)	4.33%

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CA = concurrent articulation.

The pattern of results for the ELP contrast in nonword items is similar to the results in Study 1 in which there was a significant difference in accuracy for nonword reading and auditory lexical decision, but no difference in reading latency or RT for either task.

5.0 DISCUSSION

Phonological processing is a widespread area of study despite (or perhaps because of) its abstract nature and the challenges associated with isolating the skills involved. For example, work has explored the phonological deficits of children with SLI, with particular focus on three areas: phonological memory storage capacity, phonological processing that is mediated by lexical knowledge, and core phonological processing that is not mediated by other factors (like phonological memory storage capacity, word learning, or lexical knowledge). This project was motivated in part by the limitations of current measures used to operationalize skills related to phonological processing. Nonword repetition, for example, continues to be widely used as a measure of phonological memory storage capacity, even though some have noted that it is a task that relies on multiple processes. Directly testing the phonological memory storage capacity component in nonword repetition can also have its challenges. For example, correlating or co-varying a classic phonological memory task like digit span with nonword repetition ignores other potential factors that could explain the correlation or shared variance. For instance, both tasks require an auditory perceptual component, and both tasks require spoken output. Either of these factors could create noise in isolating phonological memory storage capacity as the shared skill. To limit the potential confounds of unwanted factors contributing to the early-late phoneme effect examined in this project, a novel differential diagnosis, or process of elimination approach, was used in which tasks were selected based upon a component or skill that they *didn't* require.

If the ELP effect persists after other potential factors are eliminated, this may suggest that the core phonological processing differs for items comprised of early versus late phonemes. This may in turn indicate that differences in nonword repetition performance can emerge from differences in core phonological processing. Thus, the use of nonword repetition as a measure of phonological memory storage capacity would be called into question. Four factors were controlled or eliminated in this project and each will be discussed further below: auditory perceptual demands, articulatory demands, phonological memory capacity, and lexicality.

5.1 NONWORD REPETITION

The early-late phoneme contrast was first used by Moore and colleagues (2010) to increase the articulatory demands of a nonword repetition task. The first question in the current project addressed whether the results from Moore and colleagues could be replicated given the additional features that were controlled during stimulus development (section 3.2.3). Results showed that the ELP effect was present using the newly designed E7 and L7 nonwords at all syllable lengths and when comparing overall scores.

One difference in outcome between the two studies was performance at the 1-syllable level: in the previous work there were no significant differences in average performance between the NRT and the L8NRT for CVC nonwords, but in the current study a significant effect was observed. The failure to observe a significant effect in the prior study may have been due to the low performance levels of two participants who had 67% and 75% PPC on the NRT CVC items.

These are atypical scores when compared to the 24 (of 27) participants who scored above a 90% at the 1-syllable level on the NRT.

As reported in the design of the current stimuli, at each syllable length there were no significant durational differences between the audio recordings of E7 and L7 items. It may be noted that the actual means are numerically large between phoneme groups at the 1-syllable and 2-syllable lengths. The durational difference was approximately 50 ms between the 1-syllable average durations. In this case, the L7 items on average were longer. Although this could be seen as a potential confound, the fact that accuracy (not speed) was measured and the stimuli (only 1-syllable in length) were not over-taxing phonological storage capacity minimizes the concern for this difference. The average duration at the 2-syllable length is greater for E7 items, which works against the outcome of decreased L7 scores and, therefore, does not seem to be problematic for the current results.

Overall, the current results build upon the prior work, demonstrating that the contrast in performance between items containing early- and late-developing phonemes is robust, and not substantially affected when stricter design features are implemented. Improving the control of list differences between articulatory feature overlap, phoneme recurrence within the task, the lexicality of constituent segments, and phonotactic probability did not mute the ELP effect. In fact, large effect sizes were found in the current study as well. The results also indicated that the ELP effect increased as the number of syllables increased. Implications will be discussed below (section 5.6).

The succeeding questions aimed to address factors that potentially could be sources of the ELP effect.

5.2 ELIMINATION OF AUDITORY PERCEPTUAL DEMANDS

The second question in this project was whether the ELP effect was present after auditory perceptual demands were eliminated. After controlling the stimuli for a number of phonological and orthographic factors, results of the nonword reading task suggest that the ELP effect persisted when auditory perceptual demands were eliminated. Participants were significantly less accurate reading CVC nonwords consisting of L7 items compared to E7 CVC nonwords; however, there were no reading latency differences between phoneme types.

Further evidence that the ELP effect is robust against auditory perceptual demands comes from Study 2 in which four lexical decision tasks were administered contrasting phoneme type, presentation modality, and concurrent articulation. Study 2 was designed to examine the effects of phoneme type after minimizing articulatory demands (discussed below), but the manipulation of presentation modality – auditory lexical decision versus visual lexical decision – provides another comparison in which the auditory perceptual demands of the task have been directly manipulated. For the accuracy of nonword items, the ELP effect persisted across both auditory and visual presentations of stimuli (i.e., no phoneme type x presentation modality interaction), suggesting that auditory perceptual demands alone cannot account for the differences in the early-late phoneme contrast. It should be noted, however, that performance was near ceiling for the visual lexical decision tasks, which could have artificially eliminated the possibility of the interaction. High performance on visual items is consistent with the d' analysis, which showed a significant main effect of modality indicating that items presented visually were more discriminable than items presented aurally. However, these d' results should be interpreted cautiously since there was also a significant main effect in response bias for presentation modality. This finding of a response bias implies that participants were more prone to respond that items were words in the auditory modality, suggesting that the lexical characteristics or wordlikeness of the stimuli may have greater influence on ambiguous items presented aurally compared to visually.

Are there reasons to suspect that an interaction between phoneme type and presentation modality would be present without the potential confound of the ceiling effect? Accuracy was decreased in conditions with auditory stimuli compared to visual stimuli (collapsed across phoneme type and conditions with/without concurrent articulation). This pattern was present for both nonwords and words. To address whether there are reasons to suspect that there would be an interaction between phoneme type and presentation modality without a ceiling effect, several factors that could have contributed to the decreased performance during auditory presentations of the stimuli were examined for biases to phoneme type.

One possibility for the decreased performance with auditory stimuli is that the perceptual qualities of the stimuli affected performance at a peripheral, acoustic level of processing. It was mentioned earlier (section 2.3.2) that certain fricatives are easily confusable, such as /f-θ/ and /v-ð/, particularly when there is no verbal context (as in nonwords) and no visual support (as with audio recorded stimuli). Importantly, however, these two sets of fricatives that are considered to be highly confusable each contain an early-developing and a late-developing phoneme, so that this acoustic confusability posed no obvious disadvantage for one phoneme group over the other. The E7 items /f, v/ recur in each Study 2 list more frequently than the L7 items /θ, ð/ (with 2 to 5 more occurrences of /f, v/ than /θ, ð/ per list), which works against the results showing decreased performance on L7 items Other consonant pairs may also be susceptible to confusability errors, like /m-n/, /s-f/, and /s-z/, but these pairs again were constituents of both the early- and late-developing phonemes. In conclusion, the effect of perceptual salience on auditory stimuli does not appear to entirely account for the decreased performance on L7 items.

Another explanation for decreased performance with auditory stimuli is that participants were not performing with maximal effort. Although they did not indicate signs of fatigue or apathy during the brief testing session and had high performance levels on the visual lexical decision tasks, participants may have been rushing through tasks and, therefore, may have been more prone to errors. However, the potential effect of effort does not appear to affect one phoneme group more than the other.

A third possibility is that some of the errors reflected dialectal differences. Informally it was noted that no participant spoke using a dialect, but it is possible that experiences with different dialects (e.g., the Pittsburgh dialect) could have affected how a participant interpreted a given word or nonword in the stimulus lists. An example of the regional Pittsburgh dialect is the monophthongization of the diphthong /ɑʊ/ to /ɑ/ (e.g., 'downtown' becomes 'dahn-tahn'). Another example is the merger of /i/ and /ɪ/, so that the word 'steel' may sound similar to 'still' ("Pittsburgh English," 2012). Given these examples, it is possible that auditory presentation of a nonword like /zɪl/ could be considered a word ('zeal'). Because the dialectal differences tend to be based on vowel contrasts, there are no *a priori* predictions that this potential confound would affect the E7 and L7 items differently. However, across all four word lists in Study 2 there were five L7 items in which this type of dialectal variation could have resulted in a decision error compared to only one E7 item, creating a potential imbalance between phoneme groups.

Although the word stimuli were not the primary focus of the analyses, it was noted that the overall lexical decision accuracy for these items was decreased as well. Word items were generally controlled on a number of phonological and orthographic factors (including word frequency) in order to create a relatively balanced word-nonword environment for the nonword items in the lexical decision tasks, but other considerations like familiarity, imageability, concreteness, and semantic features were not controlled for the word stimuli. Therefore, individual items that are low frequency and/or unfamiliar to the participants (college students) could have contributed to the overall decreased scores. Additionally, there were 4 word items with bound morphemes in the E7 word lists and 6 items in the L7 lists. The overall accuracy for some of these words (e.g., ‘teed’, ‘vied’, ‘lows’) was low. The overall accuracy was 50% or less for the following E7 words in Study 2: ‘teed’, ‘vied’, and ‘fain’. The overall accuracy was 50% or less for the following L7 words in Study 2: ‘laze’, ‘lows’, and ‘thaws’. With 3 items from each phoneme group, this confound does not seem to bias one phoneme group over another, but it could account for the overall decreased accuracy observed.

Taken together, there was an ELP effect that persisted when the auditory perceptual demands were eliminated. This was apparent when participants were given visual stimuli to read in the nonword reading task. This was also apparent from the visual lexical decision tasks of Study 2, although conclusions based solely on the Study 2 tasks are susceptible to potential confounding factors.

5.3 ELIMINATION OF ARTICULATORY DEMANDS

To minimize articulatory demands, participants were asked to complete lexical decision tasks in which no spoken output was required. In Study 1, a single auditory lexical decision task was administered. In Study 2, both overt and covert articulation were considered with using a 2 x 2 x 2 experimental design that also contrasted phoneme type and presentation modality. Results suggested that the ELP effect persists when articulatory demands are minimized in the ways implemented in this project. The pattern of performance was similar in both studies: participants had significantly poorer accuracy on L7 nonwords compared to E7 nonwords, but no RT differences between phoneme types. In Study 2, performance was decreased in conditions with CA compared to those without (collapsed across phoneme type and presentation modality), suggesting that the lexical decision tasks were sensitive to processes involved with articulatory planning, inner speech rehearsal, and/or the allocation of attention across two simultaneous tasks. In nonwords, the ELP effect persisted when covert articulation was minimized (i.e., no phoneme type x CA interaction), further suggesting that the articulatory demands alone, as operationalized in this study, cannot account for the differences between phoneme type.

The size of the ELP effect was larger in the Study 1 lexical decision task than in the Study 2 lexical decision tasks. One explanation for this difference is that the stimuli in Study 1 were controlled for phonological but not orthographic factors, whereas all stimuli in Study 2 were controlled for both phonological and orthographic factors. It is possible that the difference in control parameters created a bias in favor of E7 items in Study 1. An alternative explanation is that Study 2 required more effort in light of the testing session being more difficult (e.g., having to use CA during lexical decision) which elicited an overall increase in performance that muted the ELP effect.

The same limitations addressed in the previous section could have contributed to the overall decreased accuracies (i.e., the perceptual confusability of certain phonemes, dialectal experiences in participants, their level of effort, and limitations with the word stimuli). The perceptual confusability of auditory items and level of participant effort were discussed in the previous section as having no obvious bias on one phoneme type over the other. The potential limitation of dialectal experiences could affect individual items, potentially biasing one phoneme type over another. However, the consistent findings across studies with different participant groups and different stimulus lists weaken the threat of this potential factor.

5.4 PHONOLOGICAL MEMORY STORAGE CAPACITY

Although phonological memory storage capacity was not manipulated in the current investigation, its influence was minimized or eliminated in two primary ways. First, 1-syllable CVC items were used in experimental tasks, a length well within the phonological memory storage capacity limits of typically-developing college students. Second, nonword reading and visual lexical decision were administered in which the item remained on the computer screen until the participant responded, so that phonological memory storage demands were minimal. The ELP effect persisted even when phonological memory storage demands were controlled, suggesting that the contrast between early- and late-developing phonemes is not a specific problem of phonological memory storage capacity. The evidence for a persistent ELP effect was apparent in all tasks, including the 1-syllable items in nonword repetition, nonword reading, and lexical decision.

5.5 LEXICALITY

Lexical factors were controlled by using nonwords as the focus of testing and analysis. Within each task, a number of linguistic parameters were controlled between the two phoneme groups to reduce lexical influences on the nonwords.

For the real word items in the auditory lexical decision task of Study 1, there were significant differences between phoneme types on both accuracy and RT, but there were no differences between phoneme groups in Study 2. One possible explanation for the results in Study 2 is that the processing of both E7 and L7 word items benefited more from other stored lexical information (e.g. semantics), muting any differences between phoneme groups. However, as stated previously, semantic aspects of the word stimuli were not controlled in this project and there were some word stimuli that were outliers, so conclusions about the word items should be made cautiously.

5.6 DIFFERENCES IN ACCURACY BETWEEN PHONEME GROUPS

Results from the current project demonstrate consistent performance differences in accuracy on E7 and L7 nonword items (as discussed below); however, there were no RT differences for these items between the two phoneme groups. It is possible that participants adapted to the experimental tasks by putting more emphasis on speed than accuracy. As addressed previously, another possibility is that the reading latencies and RTs are reflections of participant effort, with decreased motivation to work at an optimal level. Future work could address this concern by compensating participants based on performance.

As stated, a consistent finding across the nonword repetition, nonword reading, and lexical decision tasks of the current project was decreased accuracy for nonword items containing later-developing phonemes compared to nonword items containing early-developing phonemes. Why are the nonword items with later-developing phonemes more prone to error? Auditory perceptual, articulatory, and phonological memory storage capacity demands, as operationalized in this project, cannot entirely account for the differences between phoneme groups, based on the persistent ELP effect when these demands were not required in a task. A signal detection analysis of *d'* and β showed a pattern of results that was consistent with the accuracy results, indicating that the L7 items had decreased discriminability compared to the E7 items that cannot be explained by response bias. It is suggested here that the early-late phoneme contrast represents core phonological processing (i.e., the long-term knowledge and use of phonemes). Therefore, it is possible that the later-developing phonemes are more prone to error due to the relatively lower quality of the phonological neural networks (the phonological representations or the connections involved in retrieving the representations, as described in section 2.1).

Lower quality phonological networks could result in more errors during linguistic tasks due to the decreased ability to suppress competing information. Interactive models of speech and linguistic processing (Dell, 1986; Gupta & MacWhinney, 1997) suggest that input of a nonword results in the activation of the target item's real-word phonological neighbors to facilitate processing of the nonword. For example, when repeating or making a lexical decision on the nonword 'suzz,' input of the nonword activates the phonological information for 'buzz,' 'fuzz,' 'sub,' 'sum,' and so on. Through ongoing interactions between input information and activated representations, the 'unwanted' phonological representations, or competitors, are suppressed. However, if the later-developing phonemes have weaker neural networks, it might be more difficult to 'rule out' competitors. This competition effect results in a greater number of active representations that can influence processing of the nonword during the task and, therefore, more opportunities for error. This idea is congruent with the results from the *d'* analysis, in which the L7 word and nonword items had decreased discriminability compared to the E7 items.

In the aforementioned speech production literature, it has been suggested that these competition effects sometimes emerge during speech production planning (e.g., Dell, 1986). From this point of view, it may be surprising that there were no RT differences between E7 and L7 items. Perhaps the brevity of the CVC stimuli could account for the lack of RT differences. If competition effects are larger for L7 items during speech production planning, longer stimulus items, requiring more sequential planning, might elicit RT differences between the two phoneme groups.

Another possibility is that lower quality phonological neural networks have a slower activation rate. If participants were rushing (or trading speed for accuracy), as the lack of RT difference might suggest, then the "weaker" representations may be more prone to error.

Lower quality phonological neural networks could also lead to decreased accuracy in tasks in which phonological memory was engaged to temporarily maintain the speech sounds for short-term manipulation and retrieval (e.g., in nonword repetition and possibly the auditory lexical decision tasks). If the L7 phonological activations are relatively weak, they could be more prone to errors as they are encoded into the phonological store. Similarly, when the information is available in the phonological store for on-line retrieval and manipulation, if the quality of the L7 phonological activations is poorer, these activations might be harder to maintain in memory and more prone to decay (Baddeley, 1986; Bowey, 2006; Edwards & Lahey, 1998).

In the nonword repetition task, as the number of syllables in the nonwords increased, the ELP effect increased. This could suggest that the L7 items were more prone to error as the phonological memory storage demands were increased. Alternatively, these results from nonword repetition are consistent with the possibility that a slower activation rate of later-developing phonemes results in increased errors, since the longer syllable items need to be processed more quickly in order to provide a response within the allotted time.

The rationale for selecting the early-late phoneme manipulation was based on an idea that the phonological representations of the late-developing phonemes might have decreased centrality in neural networks and, therefore, have fewer neural connections compared to early-developing phonemes (see section 2.3.3). The current study demonstrated a difference in performance with early- and late-developing phonemes after controlling for a number of factors, suggesting that the ELP effect might represent core phonological processing. If this is true, it remains an open question as to the specific locus of difference. For example, it is possible that the quality of the phonological representations themselves is affected as certain phonemes are acquired earlier than others, or rather that the connections involved with retrieving the representations for online processing are affected. Further, some models of speech production suggest that information about the articulatory gestures used in speech is accessed at the level of the representation (e.g., Dell, 1986). If this is the case, the articulatory demands of speech sounds could result in processing differences at the representational level. The abstract nature and conceptualization of core phonological processing limits the differentiation of these constructs within core phonological processing, although making progress in this area of study would be theoretically valuable.

Perhaps these issues could be further examined by extending previous functional magnetic resonance imaging research that has investigated the neural substrates involved in the speech production network. For example, Peeva, Guenther, Tourville, Nieto-Castanon, Anton, Nazarian, and Alario (2010) reported a set of neural substrates preferentially engaged in phonemic, syllabic, and supra-syllabic levels of processing during speech. To determine if there are neural differences between early- and late-developing phonemes, future work could compare the ELP manipulation within the Peeva, et al. paradigm to determine if the locus and extent of activation varies between phoneme groups.

Another possible way to examine the underlying mechanisms involved in the ELP contrast in future work is to analyze the serial position effects in the 4-syllable nonword repetition stimuli. If the ELP effect is driven by phonological memory storage differences between E7 and L7 items, it might be expected that the magnitude of the ELP effect would increase for final syllables compared to initial syllables of the 4-syllable items. If the ELP effect is primarily driven by the difference in the quality of the phonological representations between early and late phonemes, then it might be expected that the magnitude of the ELP effect is similar across all of the syllables in the 4-syllable nonwords.

5.7 PHONOLOGICAL PROCESSING IN SLI

One of the accounts for the phonological deficits in children with SLI proposes that the locus of impairment is in the phonological memory storage capacity. It is well known that nonword repetition is a task that relies on multiple processes, but this has not precluded performance on nonword repetition from being used as a measure for phonological memory storage capacity. As recently indicated, “Indeed, the memory component of the [nonword repetition task] has perhaps been the most extensively studied, so much so that the [nonword repetition task] is generally viewed as a measure of phonological memory capacity” (Coady & Evans, 2008, p. 18).

The results from the current work demonstrate that nonword repetition performance can vary based on factors independent from phonological memory storage capacity, thus challenging the use of nonword repetition as a measure of this construct. Previous work that has challenged the use of nonword repetition as a measure solely of phonological memory storage capacity has often directly contrasted nonword repetition and other measures. This approach makes it difficult to rule out potential confounding factors that could explain the shared variance between the tasks. The current work took a differential diagnosis approach to test whether there was some underlying factor that could influence nonword repetition performance, independent of phonological memory storage capacity, auditory perceptual demands, articulatory demands, and lexical knowledge.

As described in section 2.2, Gathercole (2006) reported that the magnitude of performance differences between children with SLI and their typically-developing peers was smaller on a serial recall task than on a nonword repetition task, suggesting that another factor was contributing to the nonword repetition outcome. To supplement the phonological storage capacity account, Gathercole suggested a ‘double deficit’ account, in which phonological storage deficits in children with SLI could be exacerbated in nonword repetition because of the speed required to process the acoustic signal in nonword repetition. She suggests that typically-developing children may be better or more efficient at processing the acoustic signal or using other speech cues, like those generated by coarticulation and prosody in multisyllabic nonwords, to assist in nonword repetition performance (see also, Archibald & Gathercole, 2007).

The results from the current study could offer an alternative explanation to the ‘double deficit’ account put forth by Gathercole. The nonword repetition and auditory lexical decision tasks in the current work contained 1-syllable CVC stimuli controlled for duration, so that prosodic cues like vowel lengthening, loudness, and pitch were not readily available to facilitate performance as they might for multisyllabic items. Prosodic and coarticulatory cues were also minimized for the visual stimuli in nonword reading and visual lexical decision in which no acoustic signal was presented. Given the performance differences in these conditions in the current study, the additional factors of acoustic signal rate and prosody cannot seem to account for the early-late phoneme differences observed. Thus, the ELP effect in nonword repetition can be affected by factors other than phonological storage and the rate of the acoustic signal. This finding could be applied to children with SLI, suggesting that their performance deficits in nonword repetition also could be explained by factors other than phonological storage and acoustic signal rate, such as a core phonological processing.

Catts, Adlof, Hogan and Weismer (2005) found two distinct subgroups of children in their study of SLI. One subgroup was characterized by phonological processing and word reading deficits (comorbid with a dyslexia diagnosis). The other SLI subgroup showed deficits in semantics, syntax, and/or discourse processing apart from phonological deficits. It is possible that the double deficit account described above (Archibald & Gathercole, 2007; Gathercole, 2006) is not necessary to explain the performance of the children with specific phonological processing deficits, because their deficits in phonological processing would be pervasive regardless of additional factors like acoustic signal rate and prosodic cues. On the other hand, for the children with SLI who do not exhibit deficits in phonological processing, it is possible that their performance on nonword repetition and other similar tasks can be explained by a double deficit. That is, they may not have deficits in phonological-based tasks unless the task has a factor that additively decreases their performance.

The focus of this study was to challenge the use of the nonword repetition task as a measure of phonological storage capacity. Other work has challenged this characterization of nonword repetition tasks as well, by varying the lexicality of nonwords to assess the influence of lexical knowledge in the deficits associated with SLI and nonword repetition. For example, Metsala and Chisholm (2010) manipulated the neighborhood density of the constituent syllables in a nonword repetition task, and showed that long-term lexical knowledge contributes to performance (for a review, Coady & Evans, 2008; see also, G. Jones, et al., 2010). The potential confounds are no different in this use of the nonword repetition task than they are in its use as a measure of phonological memory storage capacity. In the current work, the early-late phoneme effect persisted after a number of phonological and orthographic factors were considered. While it is indeed reasonable that long-term lexical knowledge can influence a child's performance on nonword repetition, it cannot be ruled out that a factor like core phonological processing is the basis of the underlying deficit (which in turn may contribute to poor nonword repetition performance and lexical knowledge, as described in section 2.1).

The results of the current work indicate that the ELP effect persists after phonological memory storage capacity, lexical knowledge, auditory perceptual demands, and articulatory demands have been controlled, suggesting that the ELP effect may represent core phonological processing. These findings have been applied to the work with SLI to provide an alternative explanation for the nonword repetition deficits seen in this population. Further work could explore the hypothesis of core phonological processing deficits in children with SLI using the early-late phoneme manipulation. If children with SLI have deficits in core phonological processing, two possible patterns of performance might be predicted when compared to typically-developing peers. First, one might predict that performance on all phoneme types is impaired, so that overall performance on both E7 and L7 items is decreased in children with SLI, but the magnitude of the ELP effect is similar compared to their peers. Another prediction is that, if there is an impairment in the phonological neural networks, a “Matthew effect” (Stanovich, 1986) would occur in which the weak get weaker. That is, if the phonological representations for later-developing phonemes are relatively weak compared to those of the early-developing phonemes, they could be more greatly affected by an impaired system. In this case, the magnitude of the ELP effect could be greater for children with SLI compared to their peers.

A potential deficit in core phonological processing leaves open the question about the etiology of the deficit. As discussed in section 2.3, factors have been proposed that could explain developmental differences in the time course of phoneme acquisition (e.g., perceptual salience, articulatory demands). Regardless of originating cause, a consequence appears to be that there are differences in the core phonological representations of early versus late phonemes. In future research with children with SLI, using a battery of tasks similar to that in the current work, certain differential data patterns in the size of the ELP effect could favor one theoretical account over another.

For example, if the magnitude of the ELP effect was greater for auditory tasks compared to visual tasks, this could support the suggestion that children with SLI have perceptual deficits that lead to poor phonological processing (see also, Coady & Evans, 2008; e.g., Joanisse & Seidenberg, 1998). In section 4.4 it was noted that there was no significant interaction between phoneme type and presentation modality (auditory versus visual presentations) in lexical decision. However, a ceiling effect in the visual conditions of the task could have muted a potential interaction (section 5.8 includes a discussion of ways to address the potential ceiling effect in future work). If there is a real interaction in which the magnitude of the ELP effect is larger in auditory tasks compared to visual tasks in normal development, it might be expected that magnitude of the interaction would be even greater for some children with SLI compared to typically-developing peers. This idea is consistent with the possibility of a subgroup of children with SLI having a 'double deficit' (Archibald & Gathercole, 2007; Gathercole, 2006) in which phonological processing deficits are exacerbated by perceptual-based factors. It might be expected that this subgroup of children with SLI would have an increased magnitude of the interaction effect when compared to their peers.

In another example, the battery of tasks used in the current project could contribute to the understanding of the role of speech production in the deficits associated with SLI. As discussed in section 2.2, it has been suggested that children (or a subgroup of children) with SLI may have motor speech deficits (e.g., Archibald & Gathercole, 2006; Goffman, 2004). In the current work, a significant main effect of CA was found. If some children with SLI have deficits in speech production, it might be expected that the magnitude of the CA effect would be greater for these children compared to their typically-developing peers. Another possibility is a double deficit account. As previously discussed, Gathercole and colleagues (2006) suggested the possibility of a double deficit account related to phonological processing and perceptual-based deficits, but it is possible that a double deficit account could apply to phonological processing and subtle speech production deficits as well. In this case, the magnitude of the main effects of phoneme type and CA may not be different between children with SLI and their peers. However, if the deficit in SLI is additive, then it might be expected that the magnitude of the phoneme type x CA interaction would be greater for children with SLI compared to their peers.

A similar approach can be taken to determine the contribution of phonological memory storage capacity deficits in children with SLI. If a data pattern was observed in which the magnitude of the ELP effect was greater than that of typically-developing peers in nonword repetition but similar to that of peers in all other tasks, this could suggest that a deficit is primarily related to phonological memory storage capacity. The ELP manipulation within this battery of tasks affords many possible outcomes that might adjudicate between the possible accounts of deficits that have been associated with SLI.

5.8 STRENGTHS AND LIMITATIONS

The current work provides evidence that differences in nonword repetition cannot be taken as evidence of a specific problem in phonological memory storage capacity. Further, there are potential influences on nonword repetition performance apart from frequently studied factors like auditory perceptual demands, articulatory demands, and lexical knowledge. These results were obtained using carefully-controlled stimuli in which a number of phonological and orthographic factors were considered. Additionally, the current work implemented a 'differential diagnosis' approach to avoid some of the potential confounds that occur in correlational studies.

The manipulation of early-developing and late-developing phonemes is a novel approach that may well tap core phonological processing. Previous work has used tasks like naming, gating, and imitation to assess the quality and retrieval of phonological representations. For example, Elbro and Pallesen (2002) assessed the distinctness of phonological representations measured by the pronunciation of words in imitation. The potential confounds for this type of approach are similar to those already addressed. Specifically, conclusions about the distinctness of the phonological representation can be confounded by considerations of auditory perceptual skills, articulation, and lexical knowledge. The early-late phoneme contrast provides a potential contrast for core phonological processing apart from these confounds.

However, there were some potential limitations of the current work that could be considered in future studies. For example, the durations of the auditory stimuli were controlled in this study so that on average there were no statistically significant differences between phoneme groups; however, as previously addressed, in some instances the actual durations were seemingly large. In future work, items could be modified using sound editing software so that all stimuli are the same length. One caution to this procedure is that items that are subjected to too much shortening or lengthening may sound more synthetic to the listener, thus potentially altering the perceptual qualities of the stimuli.

As mentioned, there were many phonological and orthographic factors that were carefully controlled so that on average there were no statistically significant differences between the E7 and L7 stimuli on each factor. However, in some instances the actual values favored the E7 stimuli, particularly the phonological factors (e.g., phonotactic and biphone probabilities) in the initial phoneme position. In future work, modifying the stimuli so that the actual values of these factors are better balanced should be done cautiously so that other controls (e.g., phoneme recurrence) are not offset by the adjustments. Another approach might be to co-vary one or two factors in which the difference between the E7 and L7 items was the most pronounced.

It was noted previously that there were possible ceiling effects in performance on the visual lexical decision task in Study 2. Future work could address this by increasing the difficulty of the task. The visual items in the current work appeared and remained on the computer screen until the participant responded. An alternative approach to increase task difficulty would be to decrease the presentation time. Items were also presented clearly in white font against a black background, but the task difficulty could be altered by changing the perceptual qualities of the stimuli by blurring, dimming, or masking the items, for example.

Some items presented during the auditory lexical decision tasks could have been prone to error due to dialectal variations. This limitation is present in most studies of language and reading because it is nearly impossible to control for all of the individual experiences that participants have with languages and dialects, particularly in a university setting with college students from diverse backgrounds. To better control for this potential confound, however, one could obtain wordlikeness ratings or assess priming effects for all stimuli to determine if there were any that were more susceptible to these types of errors.

All of the L7 phonemes involve tongue movement during articulation, whereas some of the E7 items do not (e.g., /m, p, f, v/). It is possible that this articulatory difference contributed to performance differences between E7 and L7 items. Kent (1992) describes the articulatory demands of phonemes (section 2.3.2.2), but it seems that the motor demands were inferred based on the age of acquisition of the phoneme, thus potentially creating a circular argument for the contribution of articulatory demands in phoneme development. The assumption that later acquisition reflects motor demands ignores other potential factors that could contribute to phoneme acquisition, for example auditory perceptual demands or the distinctness of the phonemes within the language (i.e., how well a phoneme contrasts with others; Stokes & Surendran, 2005). Future work could examine articulatory demands in early-developing versus late-developing phonemes by using a smaller subset of phonemes that involve similar movements of the articulators. For example, in the current E7 and L7 lists, there are three E7 alveolar consonants (/n, d, t/) and three L7 alveolar consonants (/s, z, l/) that could be compared. Although these subsets of phonemes are balanced on place of articulation and voicing, they represent different manners of articulation, which could be considered in subsequent studies with varying subsets of early-developing and late-developing stimuli.

6.0 CONCLUSION

The results from this work challenge the assumption that children with SLI have deficits in phonological memory storage capacity simply because they perform poorly on nonword repetition tasks. This position was tested using a manipulation of early- versus late-developing phonemes that was predicted to show differences in nonword repetition performance even after many potential confounding factors were controlled. The results support the growing literature suggesting that nonword repetition is a task that relies on multiple processes and cannot be used as a measure of phonological memory storage capacity alone. Additionally, nonword repetition performance draws on skills apart from auditory perceptual demands, articulatory demands, and lexical knowledge. It is suggested here that core phonological processing could affect performance on nonword repetition as well as on tasks related to phonological memory storage and retrieval.

APPENDIX A

NONWORD REPETITION STIMULI

Table 11. Nonword repetition stimuli (Study 1), and mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the audio recording durations in milliseconds at each syllable length.

Syllable length	E7 stimuli	L7 stimuli	Duration in ms
One Syllable	/faɪp/	/zaɪθ/	M_{E7} (*SD*) = 709.50 (16.86)
	/vom/	/θeʃ/	M_{L7} (*SD*) = 760.75 (45.68)
	/taʊd/	/lɔɪs/	$t(6) = 2.11$
	/dɔɪf/	/raʊð/	$p = 0.08$
Two Syllable	/tævɑm/	/sæθɑʃ/	M_{E7} (*SD*) = 1158.75 (57.36)
	/maʊnɑv/	/ʃaʊzeð/	M_{L7} (*SD*) = 1082.25 (30.40)
	/vefum/	/θulaʊʃ/	$t(6) = 2.36$
	/nɔɪtef/	/ðɔɪzɑl/	$p = 0.06$
Three Syllable	/dævepɔɪn/	/θɑʃɔɪzos/	M_{E7} (*SD*) = 1405.00 (105.15)
	/mɔɪpæfaʊn/	/sɑlæður/	M_{L7} (*SD*) = 1376.75 (92.38)
	/vutaʊmæf/	/θuræðɔɪʃ/	$t(6) = 0.40$
	/faɪdɑvop/	/ʃæθɔɪraʊs/	$p = 0.70$
Four Syllable	/dɑvupɔɪtæf/	/ðusæzelɔɪθ/	M_{E7} (*SD*) = 1730.25 (74.87)
	/nɔɪfaʊmævɑt/	/rælaʊʃɔɪzɑs/	M_{L7} (*SD*) = 1743.50 (79.04)
	/tedæfumɔɪn/	/zaɪθɔɪʃæsar/	$t(6) = 0.24$
	/maʊdɔɪnɑpæv/	/ʃaʊθezolæð/	$p = 0.82$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 12. Mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the phonotactic probability of each phoneme position at each syllable length in the nonword repetition task (Study 1).

Syllable length	Phonotactic probability, position 1	Phonotactic probability, position 2	Phonotactic probability, position 3	Phonotactic probability, position 4	Phonotactic probability, position 5	Phonotactic probability, position 6	Phonotactic probability, position 7	Phonotactic probability, position 8	Phonotactic probability, position 9
1-syllable nonwords									
E7 *M* (*SD*)	0.041 (0.013)	0.024 (0.021)	0.036 (0.012)						
L7 *M* (*SD*)	0.023 (0.023)	0.020 (0.015)	0.024 (0.036)						
	$t(6) = 1.37$	$t(6) = 0.39$	$t(6) = 0.61$						
	$p = 0.22$	$p = 0.71$	$p = 0.56$						
2-syllable nonwords									
E7 *M* (*SD*)	0.037 (0.017)	0.030 (0.034)	0.051 (0.036)	0.015 (0.003)	0.020 (0.009)				
L7 *M* (*SD*)	0.030 (0.048)	0.029 (0.035)	0.030 (0.030)	0.013 (0.007)	0.019 (0.019)				
	$t(6) = 0.26$	$t(6) = 0.07$	$t(6) = 0.90$	$t(6) = 0.58$	$t(6) = 0.14$				
	$p = 0.81$	$p = 0.94$	$p = 0.41$	$p = 0.59$	$p = 0.89$				
3-syllable nonwords									
E7 *M* (*SD*)	0.045 (0.015)	0.035 (0.032)	0.041 (0.018)	0.013 (0.007)	0.018 (0.008)	0.010 (0.010)	0.049 (0.045)		
L7 *M* (*SD*)	0.031 (0.047)	0.056 (0.024)	0.042 (0.040)	0.009 (0.009)	0.015 (0.023)	0.009 (0.009)	0.047 (0.013)		
	$t(6) = 0.53$	$t(6) = -1.03$	$t(6) = -0.03$	$t(6) = 0.66$	$t(6) = 0.23$	$t(6) = 0.10$	$t(6) = 0.09$		
	$p = 0.62$	$p = 0.34$	$p = 0.98$	$p = 0.53$	$p = 0.83$	$p = 0.92$	$p = 0.93$		
4-syllable nonwords									
E7 *M* (*SD*)	0.044 (0.015)	0.026 (0.026)	0.030 (0.010)	0.008 (0.008)	0.034 (0.028)	0.011 (0.006)	0.038 (0.048)	0.005 (0.004)	0.065 (0.064)
L7 *M* (*SD*)	0.016 (0.023)	0.036 (0.030)	0.042 (0.040)	0.009 (0.008)	0.013 (0.002)	0.018 (0.013)	0.037 (0.016)	0.004 (0.003)	0.021 (0.029)
	$t(6) = 2.07$	$t(6) = -0.54$	$t(6) = -0.59$	$t(6) = -0.14$	$t(6) = 1.48$	$t(6) = -0.92$	$t(6) = 0.05$	$t(6) = 0.35$	$t(6) = 1.24$
	$p = 0.08$	$p = 0.61$	$p = 0.58$	$p = 0.89$	$p = 0.19$	$p = 0.40$	$p = 0.96$	$p = 0.74$	$p = 0.26$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 13. Mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the biphone probability of each biphone position at each syllable length in the nonword repetition task (Study 1).

Syllable length	Biphone probability, biphone 1	Biphone probability, biphone 2	Biphone probability, biphone 3	Biphone probability, biphone 4	Biphone probability, biphone 5	Biphone probability, biphone 6	Biphone probability, biphone 7	Biphone probability, biphone 8
1-syllable nonwords								
E7 *M* (*SD*)	0.0008 (0.0007)	0.0010 (0.001)						
L7 *M* (*SD*)	0.0002 (0.0002)	0.0003 (0.0003)						
	$t(6) = 1.52$	$t(6) = 1.29$						
	$p = 0.18$	$p = 0.25$						
2-syllable nonwords								
E7 *M* (*SD*)	0.0016 (0.0016)	0.0016 (0.0018)	0.0009 (0.0007)	0.0006 (0.0007)				
L7 *M* (*SD*)	0.0012 (0.0023)	0.0009 (0.0006)	0.0004 (0.0003)	0.0007 (0.0013)				
	$t(6) = 0.25$	$t(6) = 0.80$	$t(6) = 1.28$	$t(6) = 0.17$				
	$p = 0.81$	$p = 0.45$	$p = 0.25$	$p = 0.87$				
3-syllable nonwords								
E7 *M* (*SD*)	0.0010 (0.0010)	0.0018 (0.0012)	0.0006 (0.0003)	0.0002 (0.0002)	0.0004 (0.0007)	0.0004 (0.0003)		
L7 *M* (*SD*)	0.0009 (0.0011)	0.0018 (0.0028)	0.0009 (0.0010)	0.00003 (0.0001)	0.00003 (0.0001)	0.0003 (0.0005)		
	$t(6) = 0.07$	$t(6) = 0.00$	$t(6) = 0.60$	$t(6) = 1.69$	$t(6) = 1.02$	$t(6) = 0.28$		
	$p = 0.95$	$p = 1.00$	$p = 0.57$	$p = 0.14$	$p = 0.35$	$p = 0.79$		
4-syllable nonwords								
E7 *M* (*SD*)	0.0012 (0.0009)	0.0008 (0.0010)	0.0001 (0.0002)	0.0002 (0.0002)	0.0006 (0.0007)	0.0004 (0.0002)	0.0002 (0.0003)	0.0002 (0.0004)
L7 *M* (*SD*)	0.0013 (0.0025)	0.0027 (0.0040)	0.0001 (0.0001)	0.0001 (0.0002)	0.0001 (0.0002)	0.0008 (0.0006)	0.0004 (0.0006)	0.0003 (0.0005)
	$t(6) = 0.13$	$t(6) = 0.90$	$t(6) = 0.00$	$t(6) = 0.78$	$t(6) = 1.36$	$t(6) = 1.22$	$t(6) = 0.50$	$t(6) = 0.24$
	$p = 0.90$	$p = 0.41$	$p = 1.00$	$p = 0.46$	$p = 0.22$	$p = 0.27$	$p = 0.63$	$p = 0.82$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

APPENDIX B

NONWORD READING STIMULI

Table 14. Early-7 (E7) and Late-7 (L7) stimuli in the nonword reading task (Study 1).

E7 Nonwords		L7 Nonwords	
Spelling	Pronunciation	Spelling	Pronunciation
doum	/dɑʊm/	loth	/lɑθ/
daif	/def/	lish	/lɪʃ/
fape	/fep/	luthe	/luð/
foit	/fɔɪt/	rel	/rɛl/
futt	/fʌt/	riz	/rɪz/
mipe	/maɪp/	roysh	/rɔɪʃ/
meaf	/mif/	sathe	/seð/
moave	/mov/	shithe	/ʃaɪð/
nime	/naɪm/	shar	/ʃar/
neave	/niv/	sheth	/ʃɛθ/
noop	/nup/	seash	/siʃ/
poff	/pɑf/	suzz	/sʌz/
pote	/pot/	thaz	/ðæz/
pud	/pʌd/	thice	/θaɪs/
tife	/taɪf/	thouse	/ðɑʊs/
tem	/tɛm/	thear	/θir/
toove	/tuv/	thoal	/θol/
vadd	/væd/	zal	/zæl/
voum	/vɑʊm/	zoyth	/zɔɪθ/
voin	/vɔɪn/	zuss	/zʌs/

Table 15. Phoneme recurrence for the nonword reading stimulus lists (Study 1).

Phoneme recurrence	E7	L7
Consonant phoneme recurrence across task	/d/ x4	/r/ x5
	/n/ x4	/ð/ x5
	/p/ x6	/l/ x6
	/t/ x6	/s/ x6
	/v/ x6	/z/ x6
	/f/ x7	/ʃ/ x6
	/m/ x7	/θ/ x6
Vowel phoneme recurrence across task	/ɑ/ x1	/ɑ/ x2
	/æ/ x1	/æ/ x2
	/ɑɪ/ x3	/ɑɪ/ x2
	/ɑʊ/ x2	/ɑʊ/ x1
	/e/ x2	/e/ x1
	/ɛ/ x1	/ɛ/ x2
	/i/ x2	/i/ x2
	/ɪ/ x0	/ɪ/ x2
	/o/ x2	/o/ x1
	/ɔɪ/ x2	/ɔɪ/ x2
	/u/ x2	/u/ x1
	/ʌ/ x2	/ʌ/ x2

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 16. Mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the orthographic factors controlled in the nonword reading task (Study 1).

Phoneme type	Number of letters	Orthographic neighborhood	Mean bigram frequency	Summed bigram frequency by position
E7 *M* (*SD*)	4.050 (0.510)	7.100 (4.876)	2426.40 (1296.48)	1514.90 (520.38)
L7 *M* (*SD*)	4.500 (0.889)	5.050 (5.125)	2971.04 (1240.65)	1921.05 (832.73)
	$t(38) = 1.96$	$t(38) = 1.30$	$t(38) = 1.36$	$t(38) = 1.85$
	$p = 0.06$	$p = 0.20$	$p = 0.18$	$p = 0.07$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 17. Mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the phonological factors controlled in the nonword reading task (Study 1).

Phoneme type	Phonological neighborhood	Weighted phonological neighborhood	Phonotactic probability, position 1	Phonotactic probability, position 2	Phonotactic probability, position 3	Biphone probability, biphone 1	Biphone probability, biphone 2
E7 *M* (*SD*)	16.70 (5.89)	3330.8 (6341.6)	0.047 (0.021)	0.034 (0.021)	0.041 (0.020)	0.0024 (0.0020)	0.0013 (0.0013)
L7 *M* (*SD*)	16.15 (11.31)	3474.6 (6439.5)	0.031 (0.035)	0.047 (0.029)	0.036 (0.034)	0.0025 (0.0042)	0.0027 (0.0042)
	$t(38) = 0.19$	$t(38) = 0.07$	$t(38) = 1.74$	$t(38) = 1.64$	$t(38) = 0.56$	$t(38) = 0.08$	$t(38) = 1.34$
	$p = 0.85$	$p = 0.94$	$p = 0.09$	$p = 0.11$	$p = 0.58$	$p = 0.94$	$p = 0.19$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

APPENDIX C

AUDITORY LEXICAL DECISION STIMULI

Table 18. Word and nonword stimuli for the auditory lexical decision task in Study 1.

Early-7 List		Late-7 List	
Nonwords	Words	Nonwords	Words
/don/	/din/	/laɪʃ/	/liʃ/
/daɪt/	/daɪv/	/lɛʃ/	/lɛs/
/fæp/	/fæn/	/loθ/	/lor/
/fiv/	/fit/	/rus/	/rɛs/
/fɑm/	/fom/	/roʃ/	/roz/
/mɪv/	/mɪt/	/ʃul/	/ʃiθ/
/mup/	/mɑp/	/ʃis/	/ʃol/
/nef/	/nep/	/sɔɪθ/	/sɪl/
/nɑʊd/	/nod/	/soð/	/suð/
/pɛm/	/pɛt/	/θir/	/θim/*
/pɑʊf/	/puf/	/θɑʊz/	/θaɪz/
/toɪm/	/tim/	/ðaɪs/	/ðaɪn/*
/tɑv/	/tɑp/	/ðez/	/ðoz/
/vaɪf/	/vaɪn/	/zɛl/	/zil/
/vid/	/voɪd/	/zuθ/	/zum/*

* Indicates 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 19. Phoneme recurrence for the auditory lexical decision word and nonword lists (Study 1).

	Nonword list		Word list	
Phoneme recurrence	E7	L7	E7	L7*
Consonant phoneme recurrence across list	/d/ x4	/l/ x5	/d/ x4	/l/ x6
	/f/ x6	/r/ x3	/f/ x4	/r/ x3
	/m/ x5	/s/ x5	/m/ x4	/s/ x4
	/n/ x3	/ʃ/ x5	/n/ x5	/ʃ/ x3
	/p/ x4	/θ/ x5	/p/ x5	/θ/ x3
	/t/ x3	/ð/ x3	/t/ x5	/ð/ x3
	/v/ x5	/z/ x4	/v/ x3	/z/ x5
				/m/ x2
				/n/ x1
Vowel phoneme recurrence across list	/ɑ/ x2	/ɑ/ x0	/ɑ/ x2	/ɑ/ x0
	/æ/ x1	/æ/ x0	/æ/ x1	/æ/ x0
	/ɑɪ/ x2	/ɑɪ/ x2	/ɑɪ/ x2	/ɑɪ/ x2
	/ɑʊ/ x2	/ɑʊ/ x1	/ɑʊ/ x0	/ɑʊ/ x0
	/e/ x1	/e/ x1	/e/ x1	/e/ x1
	/ɛ/ x1	/ɛ/ x2	/ɛ/ x1	/ɛ/ x1
	/i/ x2	/i/ x2	/i/ x3	/i/ x4
	/ɪ/ x1	/ɪ/ x0	/ɪ/ x1	/ɪ/ x1
	/o/ x1	/o/ x3	/o/ x2	/o/ x4
	/ɔɪ/ x1	/ɔɪ/ x1	/ɔɪ/ x1	/ɔɪ/ x0
	/u/ x1	/u/ x3	/u/ x1	/u/ x2

* Includes the phoneme counts for three 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were included in the phoneme recurrence count, but were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 20. Mean (*M*), standard deviation (*SD*), independent sample *t*-test (*t*), and *p*-value (*p*) for the phonological factors controlled in the auditory lexical decision task (Study 1).

Stimuli	Phonological neighborhood	Weighted phonological neighborhood	Phonotactic probability, position 1	Phonotactic probability, position 2	Phonotactic probability, position 3	Biphone probability, biphone 1	Biphone probability, biphone 2
Nonwords							
E7 *M* (*SD*)	17.73 (10.68)	1338.87 (1070.06)	0.047 (0.020)	0.042 (0.027)	0.039 (0.021)	0.0023 (0.0021)	0.0021 (0.0019)
L7 *M* (*SD*)	17.20 (8.57)	1977.80 (2042.80)	0.030 (0.034)	0.036 (0.020)	0.037 (0.034)	0.0013 (0.0014)	0.0012 (0.0022)
	$t(28) = 0.15$	$t(28) = 1.07$	$t(28) = 1.69$	$t(28) = 0.69$	$t(28) = 0.25$	$t(28) = 1.42$	$t(28) = 1.22$
	$p = 0.88$	$p = 0.29$	$p = 0.10$	$p = 0.50$	$p = 0.80$	$p = 0.17$	$p = 0.23$
Words							
E7 *M* (*SD*)	26.73 (12.14)	2663.53 (3251.64)	0.047 (0.020)	0.046 (0.024)	0.054 (0.026)	0.0032 (0.0023)	0.0033 (0.0035)
L7 *M* (*SD*)	30.50 (15.86)	1748.67 (1213.60)	0.037 (0.035)	0.046 (0.021)	0.045 (0.033)	0.0024 (0.0026)	0.0041 (0.0052)
	$t(25) = 0.70$	$t(25) = 0.92$	$t(25) = 0.99$	$t(25) = 0.02$	$t(25) = 0.86$	$t(25) = 0.90$	$t(25) = 0.50$
	$p = 0.49$	$p = 0.37$	$p = 0.33$	$p = 0.98$	$p = 0.40$	$p = 0.38$	$p = .62$

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

APPENDIX D

STUDY 2 LEXICAL DECISION STIMULI

Table 21. Word and nonword stimuli for List 1 of the lexical decision task in Study 2.

Stimulus Lists	Nonwords Spelling	Nonwords Pronunciation	Words Spelling	Words Pronunciation
Early-7 list	daif	/def/	dawn	/dɔn/
	dutt	/dʌt/	deep	/dip/
	foat	/fot/	fade	/fed/
	fute	/fut/	five	/faɪv/
	meef	/mif/	fought	/fɔt/
	mipe	/maɪp/	knife	/naɪf/
	neave	/niv/	mauve	/mɔv/
	nime	/naɪm/	moon	/mun/
	paim	/pem/	note	/not/
	pote	/pot/	pout	/pɑʊt/
	teeve	/tiv/	putt	/pʌt/
	toff	/tɔf/	teed	/tid/
	vapp	/væp/	town	/tɑʊn/
	vife	/vaɪf/	vied	/vaɪd/
	vean	/vin/	van	/væn/
Late-7 list	laz	/læz/	laws	/lɔz/
	saith	/seθ/	laze	/lez/
	lesh	/lɛʃ/	loathe	/loð/
	rel	/rɛl/	race	/res/
	reez	/riz/	rise	/raɪz/
	shuss	/ʃʌs/	rush	/rʌʃ/
	shule	/ʃul/	sash	/sæʃ/
	sosh	/sɔʃ/	sear	/sir/
	sithe	/saɪð/	sell	/sɛl/
	thace	/ðes/	shawl	/ʃɔl/
	thauz	/θɑʊz/	shoes	/ʃuz/
	thole	/ðol/	this	/ðɪs/
	thoss	/θɔs/	thin*	/θɪn/*
	zal	/zæl/	thought*	/θɔt/*
	zole	/zol/	zoom*	/zum/*

* Indicates 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 22. Word and nonword stimuli for List 2 of the lexical decision task in Study 2.

Stimulus Lists	Nonwords Spelling	Nonwords Pronunciation	Words Spelling	Words Pronunciation
Early-7 list	dight	/daɪt/	date	/det/
	dute	/dut/	dine	/daɪn/
	fape	/fep/	fame	/fem/
	foon	/fun/	feed	/fid/
	mide	/maɪd/	fight	/faɪt/
	tep	/tɛp/	knot	/nɑt/
	veem	/vim/	mean	/min/
	naid	/ned/	mood	/mud/
	noop	/nup/	node	/nod/
	pime	/paɪm/	peeve	/piv/
	peaf	/pif/	pain	/pen/
	taff	/tæf/	tune	/tun/
	toove	/tuv/	type	/taɪp/
	vade	/ved/	vet	/vɛt/
	vome	/vom/	void	/vɔɪd/
Late-7 list	zool	/zul/	leash	/liʃ/
	loath	/loθ/	lies	/laɪz/
	luss	/lʌs/	loss	/lɔs/
	rall	/rɔl/	rail	/rel/
	suzz	/sʌz/	rose	/roz/
	saze	/sez/	seal	/sil/
	seash	/siʃ/	shell	/ʃɛl/
	shez	/ʃɛz/	size	/saɪz/
	shile	/ʃaɪl/	soar	/sor/
	thar	/θɑr/	sure	/ʃur/
	thil	/θɪl/	these	/ðiz/
	thure	/ður/	wrath	/ræθ/
	thush	/ðʌʃ/	theme*	/θim/*
	zel	/zɛl/	thud*	/θʌd/*
	soth	/sɔθ/	zone*	/zon/*

* Indicates 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 23. Word and nonword stimuli for List 3 of the lexical decision task in Study 2.

Stimulus	Nonwords		Words	
Lists	Spelling	Pronunciation	Spelling	Pronunciation
Early-7 list	deave	/div/	dime	/daɪm/
	doif	/dɔɪf/	doubt	/dɑʊt/
	feam	/fim/	fain	/fen/
	feeve	/fiv/	foam	/fom/
	vune	/vun/	foot	/fʊt/
	mafe	/mef/	knight	/naɪt/
	mupe	/mup/	main	/men/
	nipe	/naɪp/	moat	/mot/
	noove	/nuv/	need	/nid/
	pode	/pod/	pave	/pev/
	poun	/pɑʊn/	poof	/puf/
	teep	/tip/	team	/tim/
	tem	/tɛm/	tied	/taɪd/
	vight	/vaɪt/	vat	/væt/
	ved	/vɛd/	vowed	/vɑʊd/
Late-7 list	soith	/sɔɪθ/	lice	/laɪs/
	luth	/luθ/	lows	/loz/
	losh	/lɑʃ/	lure	/lur/
	rith	/rɪθ/	raise	/rez/
	ruzz	/rʌz/	rash	/ræʃ/
	sar	/sɑr/	role	/rol/
	shoss	/ʃɔs/	shale	/ʃel/
	shiz	/ʃɪz/	shore	/ʃor/
	sile	/saɪl/	sill	/sɪl/
	thale	/ðel/	south	/sɑʊθ/
	thall	/θɔl/	those	/ðoz/
	thice	/θaɪs/	zeal	/zil/
	thel	/ðɛl/	that*	/ðæt/*
	zoice	/zɔɪs/	thumb*	/θʌm/*
	zus	/zʌs/	zip*	/zɪp/*

* Indicates 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 24. Word and nonword stimuli for List 4 of the lexical decision task in Study 2.

Stimulus Lists	Nonwords Spelling	Nonwords Pronunciation	Words Spelling	Words Pronunciation
Early-7 list	dipe	/daɪp/	dean	/din/
	doan	/don/	dive	/daɪv/
	fode	/fod/	fate	/fet/
	fett	/fɛt/	food	/fud/
	meave	/miv/	knit	/nɪt/
	mep	/mɛp/	might	/maɪt/
	neam	/nim/	mode	/mod/
	noit	/nɔɪt/	neat	/nit/
	peff	/pɛf/	paid	/ped/
	poom	/pum/	phone	/fon/
	taid	/ted/	pine	/paɪn/
	tife	/taɪf/	time	/taɪm/
	vadd	/væd/	towed	/tod/
	vate	/vet/	vain	/ven/
	veet	/vit/	vote	/vot/
Late-7 list	leeth	/liθ/	lace	/les/
	sish	/sɪʃ/	lash	/læʃ/
	luthe	/luð/	lore	/lor/
	ral	/ræl/	lose	/luz/
	riz	/rɪz/	rice	/raɪs/
	soush	/sɑʊʃ/	rule	/rul/
	shar	/ʃɑr/	sale	/sel/
	shoth	/ʃɔθ/	sews	/soz/
	sule	/sul/	share	/ʃɛr/
	sazz	/sæz/	shies	/ʃaɪz/
	thaze	/ðez/	thaws	/θɔz/
	theal	/ðil/	thus	/ðʌs/
	thosh	/θɔʃ/	them*	/ðɛm/*
	zoil	/zɔɪl/	thief*	/θif/*
	zil	/zɪl/	zap*	/zæp/*

* Indicates 'mixed' words that contain an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word. These words were used due to the limited possibilities of real CVC words that consist of only L7 phonemes and that meet all of the other criteria. These 'mixed' words were not included in analyses with E7 and L7 comparisons.

Note. E7 = Early-7 stimuli; L7 = Late-7 stimuli; CVC = consonant-vowel-consonant.

Table 25. Phoneme recurrence comparing Early-7 (E7) and Late-7 (L7) stimuli in the Study 2 lexical decision nonword lists.

	Nonword list 1		Nonword list 2		Nonword list 3		Nonword list 4	
Phoneme recurrence	E7	L7	E7	L7	E7	L7	E7	L7
Consonant phoneme recurrence across list	/d/ x2	/l/ x7	/d/ x5	/l/ x7	/d/ x4	/l/ x7	/d/ x5	/l/ x6
	/f/ x6	/r/ x2	/f/ x4	/r/ x4	/f/ x4	/r/ x6	/f/ x4	/r/ x3
	/m/ x4	/s/ x6	/m/ x4	/s/ x5	/m/ x4	/s/ x3	/m/ x4	/s/ x4
	/n/ x3	/ʃ/ x4	/n/ x3	/ʃ/ x4	/n/ x4	/ʃ/ x3	/n/ x3	/ʃ/ x5
	/p/ x4	/θ/ x3	/p/ x5	/θ/ x4	/p/ x5	/θ/ x5	/p/ x4	/θ/ x3
	/t/ x6	/ð/ x3	/t/ x5	/ð/ x2	/t/ x3	/ð/ x2	/t/ x6	/ð/ x3
	/v/ x5	/z/ x5	/v/ x4	/z/ x4	/v/ x6	/z/ x4	/v/ x4	/z/ x5
Vowel phoneme recurrence across list	/ɑ/ x0	/ɑ/ x0	/ɑ/ x0	/ɑ/ x1	/ɑ/ x0	/ɑ/ x2	/ɑ/ x0	/ɑ/ x1
	/æ/ x1	/æ/ x2	/æ/ x1	/æ/ x0	/æ/ x0	/æ/ x0	/æ/ x1	/æ/ x2
	/ɑɪ/ x3	/ɑɪ/ x1	/ɑɪ/ x3	/ɑɪ/ x1	/ɑɪ/ x2	/ɑɪ/ x2	/ɑɪ/ x2	/ɑɪ/ x0
	/ɑʊ/ x0	/ɑʊ/ x1	/ɑʊ/ x0	/ɑʊ/ x0	/ɑʊ/ x1	/ɑʊ/ x0	/ɑʊ/ x0	/ɑʊ/ x1
	/e/ x2	/e/ x2	/e/ x3	/e/ x1	/e/ x1	/e/ x1	/e/ x2	/e/ x1
	/ɛ/ x0	/ɛ/ x2	/ɛ/ x1	/ɛ/ x2	/ɛ/ x2	/ɛ/ x1	/ɛ/ x3	/ɛ/ x0
	/i/ x4	/i/ x1	/i/ x2	/i/ x1	/i/ x4	/i/ x0	/i/ x3	/i/ x2
	/ɪ/ x0	/ɪ/ x0	/ɪ/ x0	/ɪ/ x1	/ɪ/ x0	/ɪ/ x2	/ɪ/ x0	/ɪ/ x3
	/o/ x2	/o/ x2	/o/ x1	/o/ x1	/o/ x1	/o/ x0	/o/ x2	/o/ x0
	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x1	/ɔɪ/ x2	/ɔɪ/ x1	/ɔɪ/ x1
	/u/ x1	/u/ x1	/u/ x4	/u/ x2	/u/ x3	/u/ x1	/u/ x1	/u/ x2
	/ʌ/ x1	/ʌ/ x1	/ʌ/ x0	/ʌ/ x3	/ʌ/ x0	/ʌ/ x2	/ʌ/ x0	/ʌ/ x0
	/ɔ/ x1	/ɔ/ x2	/ɔ/ x0	/ɔ/ x2	/ɔ/ x0	/ɔ/ x2	/ɔ/ x0	/ɔ/ x2

Table 26. Phoneme recurrence comparing Early-7 (E7) and Late-7 (L7) stimuli in the Study 2 lexical decision word lists.

	Word list 1		Word list 2		Word list 3		Word list 4	
Phoneme recurrence	E7	L7	E7	L7	E7	L7	E7	L7
Consonant phoneme recurrence across list	/d/ x5	/l/ x5	/d/ x6	/l/ x6	/d/ x5	/l/ x7	/d/ x6	/l/ x6
	/f/ x4	/r/ x4	/f/ x3	/r/ x5	/f/ x4	/r/ x5	/f/ x3	/r/ x4
	/m/ x2	/s/ x5	/m/ x3	/s/ x4	/m/ x5	/s/ x3	/m/ x3	/s/ x5
	/n/ x6	/ʃ/ x4	/n/ x6	/ʃ/ x3	/n/ x4	/ʃ/ x3	/n/ x6	/ʃ/ x3
	/p/ x3	/θ/ x2	/p/ x3	/θ/ x3	/p/ x2	/θ/ x2	/p/ x2	/θ/ x2
	/t/ x6	/ð/ x2	/t/ x6	/ð/ x1	/t/ x7	/ð/ x2	/t/ x7	/ð/ x2
	/v/ x4	/z/ x5	/v/ x3	/z/ x5	/v/ x3	/z/ x4	/v/ x3	/z/ x4
		/m/ x1		/d/ x1		/m/ x1		/f/ x1
		/n/ x1		/m/ x1		/p/ x1		/m/ x1
		/t/ x1		/n/ x1		/t/ x1		/p/ x1
Vowel phoneme recurrence across list	/ɑ/ x0	/ɑ/ x0	/ɑ/ x1	/ɑ/ x0	/ɑ/ x0	/ɑ/ x0	/ɑ/ x0	/ɑ/ x0
	/æ/ x1	/æ/ x1	/æ/ x0	/æ/ x1	/æ/ x1	/æ/ x2	/æ/ x0	/æ/ x2
	/ɑɪ/ x3	/ɑɪ/ x1	/ɑɪ/ x3	/ɑɪ/ x2	/ɑɪ/ x3	/ɑɪ/ x1	/ɑɪ/ x4	/ɑɪ/ x2
	/ɑʊ/ x2	/ɑʊ/ x0	/ɑʊ/ x0	/ɑʊ/ x0	/ɑʊ/ x2	/ɑʊ/ x1	/ɑʊ/ x0	/ɑʊ/ x0
	/e/ x1	/e/ x2	/e/ x3	/e/ x1	/e/ x3	/e/ x2	/e/ x3	/e/ x2
	/ɛ/ x0	/ɛ/ x1	/ɛ/ x1	/ɛ/ x1	/ɛ/ x0	/ɛ/ x0	/ɛ/ x0	/ɛ/ x2
	/i/ x2	/i/ x1	/i/ x3	/i/ x4	/i/ x2	/i/ x1	/i/ x2	/i/ x1
	/ɪ/ x0	/ɪ/ x2	/ɪ/ x0	/ɪ/ x0	/ɪ/ x0	/ɪ/ x2	/ɪ/ x1	/ɪ/ x0
	/o/ x1	/o/ x1	/o/ x1	/o/ x3	/o/ x2	/o/ x4	/o/ x4	/o/ x2
	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x1	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x0	/ɔɪ/ x0
	/ʊ/ x0	/ʊ/ x0	/ʊ/ x0	/ʊ/ x0	/ʊ/ x1	/ʊ/ x0	/ʊ/ x0	/ʊ/ x0
	/u/ x1	/u/ x2	/u/ x2	/u/ x1	/u/ x1	/u/ x1	/u/ x1	/u/ x2
	/ʌ/ x1	/ʌ/ x1	/ʌ/ x0	/ʌ/ x1	/ʌ/ x0	/ʌ/ x1	/ʌ/ x0	/ʌ/ x1
	/ɔ/ x3	/ɔ/ x3	/ɔ/ x0	/ɔ/ x1	/ɔ/ x0	/ɔ/ x0	/ɔ/ x0	/ɔ/ x1

Table 27. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the phonological factors controlled for the nonwords in the four Study 2 stimulus lists.

Statistical effects	Phonological neighborhood	Weighted phonological neighborhood	Phonotactic probability, position 1	Phonotactic probability, position 2	Phonotactic probability, position 3	Biphone probability, biphone 1	Biphone probability, biphone 2
E7 Nonwords							
M_1 (*SD*)	19.47 (7.34)	1388.00 (778.47)	0.05 (0.02)	0.04 (0.01)	0.04 (0.02)	0.002 (0.002)	0.002 (0.001)
M_2 (*SD*)	22.67 (9.90)	3597.93 (7120.78)	0.04 (0.02)	0.04 (0.02)	0.04 (0.02)	0.002 (0.001)	0.002 (0.001)
M_3 (*SD*)	19.93 (7.40)	1808.40 (1374.91)	0.05 (0.02)	0.03 (0.02)	0.04 (0.02)	0.002 (0.002)	0.002 (0.001)
M_4 (*SD*)	23.00 (9.30)	2208.53 (1690.20)	0.05 (0.02)	0.04 (0.02)	0.05 (0.02)	0.002 (0.002)	0.002 (0.001)
L7 Nonwords							
M_1 (*SD*)	20.33 (13.38)	2443.13 (3259.45)	0.03 (0.04)	0.04 (0.02)	0.05 (0.03)	0.002 (0.002)	0.003 (0.003)
M_2 (*SD*)	21.00 (12.35)	6870.20 (17637.96)	0.04 (0.04)	0.04 (0.02)	0.05 (0.03)	0.001 (0.002)	0.003 (0.005)
M_3 (*SD*)	19.27 (11.03)	3210.87 (4883.88)	0.03 (0.04)	0.04 (0.03)	0.05 (0.03)	0.002 (0.004)	0.003 (0.004)
M_4 (*SD*)	21.73 (11.28)	3442.53 (5014.75)	0.04 (0.04)	0.05 (0.03)	0.04 (0.03)	0.003 (0.005)	0.003 (0.005)
$ME_{PhonemeType}$	$F(1) = 0.13$	$F(1) = 1.70$	$F(1) = 1.99$	$F(1) = 1.84$	$F(1) = 0.00$	$F(1) = 0.11$	$F(1) = 3.88$
	$p = 0.72$	$p = 0.19$	$p = 0.16$	$p = 0.18$	$p = 0.99$	$p = 0.74$	$p = 0.05$
ME_{List}	$F(3) = 0.52$	$F(3) = 1.20$	$F(3) = 0.07$	$F(3) = 0.48$	$F(3) = 0.17$	$F(3) = 0.66$	$F(3) = 0.03$
	$p = 0.67$	$p = 0.32$	$p = 0.98$	$p = 0.70$	$p = 0.92$	$p = 0.58$	$p = 0.99$
PhonemeType x List	$F(3) = 0.09$	$F(3) = 0.15$	$F(3) = 0.08$	$F(3) = 0.06$	$F(3) = 0.54$	$F(3) = 0.42$	$F(3) = 0.07$
	$p = 0.97$	$p = 0.93$	$p = 0.97$	$p = 0.98$	$p = 0.65$	$p = 0.74$	$p = 0.98$

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

Table 28. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the orthographic factors controlled for the nonwords in the four Study 2 stimulus lists.

Statistical effects	Number of letters	Orthographic neighborhood	Mean bigram frequency	Summed bigram frequency by position	Number of Friends	Consistency Ratio
E7 Nonwords						
M_1 (*SD*)	4.13 (0.35)	7.67 (4.29)	2549.31 (1179.39)	1634.20 (608.64)	7.13 (4.09)	1.00 (0.00)
M_2 (*SD*)	4.00 (0.53)	9.73 (4.57)	2564.44 (921.67)	1627.87 (514.81)	10.93 (5.70)	0.95 (0.09)
M_3 (*SD*)	4.13 (0.64)	8.20 (4.49)	2681.55 (1380.20)	1629.67 (749.36)	8.07 (5.68)	1.00 (0.00)
M_4 (*SD*)	4.00 (0.38)	8.27 (4.70)	2688.41 (1115.63)	1658.20 (600.52)	8.47 (5.45)	0.99 (0.06)
L7 Nonwords						
M_1 (*SD*)	4.33 (0.82)	5.93 (4.95)	3273.63 (1499.09)	1945.33 (643.33)	6.47 (5.42)	0.98 (0.06)
M_2 (*SD*)	4.27 (0.59)	6.93 (5.01)	2683.81 (1200.94)	1806.67 (703.13)	7.67 (6.04)	0.94 (0.11)
M_3 (*SD*)	4.27 (0.70)	7.80 (6.67)	2946.24 (1281.55)	1736.13 (721.53)	9.40 (6.88)	0.98 (0.06)
M_4 (*SD*)	4.27 (0.80)	6.20 (5.62)	2875.92 (1322.19)	1734.00 (706.95)	6.60 (5.67)	0.95 (0.11)
$ME_{PhonemeType}$	$F(1) = 3.62$	$F(1) = 3.55$	$F(1) = 2.02$	$F(1) = 1.94$	$F(1) = 1.17$	$F(1) = 3.06$
	$p = 0.06$	$p = 0.06$	$p = 0.16$	$p = 0.17$	$p = 0.28$	$p = 0.08$
ME_{List}	$F(3) = 0.19$	$F(3) = 0.57$	$F(3) = 0.27$	$F(3) = 0.16$	$F(3) = 1.20$	$F(3) = 2.32$
	$p = 0.90$	$p = 0.64$	$p = 0.84$	$p = 0.93$	$p = 0.31$	$p = 0.08$
PhonemeType x List	$F(3) = 0.08$	$F(3) = 0.29$	$F(3) = 0.36$	$F(3) = 0.19$	$F(3) = 0.89$	$F(3) = 0.05$
	$p = 0.97$	$p = 0.83$	$p = 0.78$	$p = 0.90$	$p = 0.45$	$p = 0.98$

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

Table 29. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the phonological factors controlled for the words in the four Study 2 stimulus lists.

Statistical effects	Phonological neighborhood	Weighted phonological neighborhood	Phonotactic probability, position 1	Phonotactic probability, position 2	Phonotactic probability, position 3	Biphone probability, biphone 1	Biphone probability, biphone 2
E7 Words							
M_1 (*SD*)	24.93 (9.02)	3848.87 (3474.78)	0.05 (0.02)	0.03 (0.02)	0.06 (0.03)	0.002 (0.001)	0.003 (0.003)
M_2 (*SD*)	31.07 (10.66)	4295.40 (7201.79)	0.05 (0.02)	0.03 (0.02)	0.06 (0.03)	0.002 (0.001)	0.002 (0.001)
M_3 (*SD*)	26.20 (13.84)	3401.13 (4221.63)	0.05 (0.02)	0.03 (0.02)	0.06 (0.02)	0.002 (0.002)	0.002 (0.002)
M_4 (*SD*)	32.33 (8.98)	2906.00 (3655.14)	0.05 (0.02)	0.04 (0.02)	0.06 (0.02)	0.002 (0.001)	0.003 (0.001)
L7 Words							
M_1 (*SD*)	29.50 (14.79)	1706.67 (1514.85)	0.05 (0.04)	0.04 (0.03)	0.04 (0.03)	0.003 (0.002)	0.003 (0.005)
M_2 (*SD*)	34.92 (18.47)	1870.83 (1169.88)	0.05 (0.04)	0.04 (0.02)	0.04 (0.03)	0.002 (0.001)	0.003 (0.005)
M_3 (*SD*)	34.08 (17.65)	1701.17 (1386.54)	0.04 (0.03)	0.04 (0.02)	0.05 (0.03)	0.003 (0.003)	0.004 (0.005)
M_4 (*SD*)	30.83 (13.86)	2421.17 (2326.95)	0.04 (0.03)	0.04 (0.02)	0.05 (0.03)	0.002 (0.001)	0.004 (0.005)
$ME_{PhonemeType}$	$F(1) = 2.00$	$F(1) = 5.23$	$F(1) = 0.34$	$F(1) = 3.80$	$F(1) = 4.58$	$F(1) = 0.39$	$F(1) = 1.70$
	$p = 0.16$	$p = 0.02^*$	$p = 0.56$	$p = 0.05$	$p = 0.04^*$	$p = 0.54$	$p = 0.20$
ME_{List}	$F(3) = 0.89$	$F(3) = 0.10$	$F(3) = 0.23$	$F(3) = 0.14$	$F(3) = 0.40$	$F(3) = 0.21$	$F(3) = 0.04$
	$p = 0.45$	$p = 0.96$	$p = 0.88$	$p = 0.93$	$p = 0.75$	$p = 0.89$	$p = 0.99$
PhonemeType x List	$F(3) = 0.55$	$F(3) = 0.34$	$F(3) = 0.23$	$F(3) = 0.63$	$F(3) = 0.26$	$F(3) = 1.09$	$F(3) = 0.14$
	$p = 0.65$	$p = 0.80$	$p = 0.88$	$p = 0.60$	$p = 0.86$	$p = 0.36$	$p = 0.93$

* Significant at $p < 0.05$ level.

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

Table 30. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the orthographic factors controlled for the words in the four Study 2 stimulus lists.

Statistical effects	Number of letters	Orthographic neighborhood	Mean bigram frequency	Summed bigram frequency by position	Number of Friends	Consistency Ratio
E7 Words						
M_1 (*SD*)	4.20 (0.68)	9.33 (5.27)	2750.89 (1522.69)	1658.53 (610.26)	12.07 (6.05)	0.94 (0.13)
M_2 (*SD*)	4.07 (0.46)	11.07 (6.66)	3217.27 (2087.79)	1720.33 (805.87)	13.67 (8.58)	0.95 (0.21)
M_3 (*SD*)	4.20 (0.68)	9.80 (5.20)	3490.26 (2015.56)	1729.87 (529.38)	13.40 (9.00)	0.87 (0.26)
M_4 (*SD*)	4.20 (0.41)	11.60 (5.34)	3884.36 (1787.24)	1993.47 (733.51)	16.00 (8.07)	0.87 (0.23)
L7 Words						
M_1 (*SD*)	4.33 (0.65)	11.33 (6.92)	3235.04 (1233.19)	1977.08 (745.94)	12.33 (8.43)	0.78 (0.28)
M_2 (*SD*)	4.33 (0.49)	10.17 (6.89)	4004.40 (1460.55)	1990.83 (676.66)	9.17 (8.08)	0.90 (0.23)
M_3 (*SD*)	4.42 (0.51)	11.17 (5.61)	3639.29 (982.70)	2144.50 (561.33)	15.50 (10.04)	0.87 (0.18)
M_4 (*SD*)	4.25 (0.45)	13.83 (6.49)	3592.93 (1410.94)	1920.50 (761.73)	14.67 (10.04)	0.82 (0.34)
$ME_{PhonemeType}$	$F(1) = 2.42$	$F(1) = 1.01$	$F(1) = 0.79$	$F(1) = 3.09$	$F(1) = 1.24$	$F(1) = 1.83$
	$p = 0.12$	$p = 0.32$	$p = 0.38$	$p = 0.08$	$p = 0.30$	$p = 0.18$
ME_{List}	$F(3) = 0.20$	$F(3) = 0.93$	$F(3) = 1.09$	$F(3) = 0.25$	$F(3) = 0.27$	$F(3) = 0.56$
	$p = 0.90$	$p = 0.43$	$p = 0.36$	$p = 0.86$	$p = 0.60$	$p = 0.64$
PhonemeType x List	$F(3) = 0.20$	$F(3) = 0.37$	$F(3) = 0.53$	$F(3) = 0.64$	$F(3) = 0.72$	$F(3) = 0.53$
	$p = 0.90$	$p = 0.77$	$p = 0.66$	$p = 0.59$	$p = 0.55$	$p = 0.67$

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

Table 31. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the auditory stimuli duration for the Study 2 stimulus lists.

Statistical Effects	Duration of Recorded Auditory Stimuli (ms)
E7 Nonwords	
M_1 (*SD*)	739.53 (53.12)
M_2 (*SD*)	732.87 (56.26)
M_3 (*SD*)	748.27 (48.03)
M_4 (*SD*)	744.80 (56.20)
L7 Nonwords	
M_1 (*SD*)	758.40 (51.87)
M_2 (*SD*)	738.73 (32.89)
M_3 (*SD*)	762.73 (34.10)
M_4 (*SD*)	765.13 (26.46)
$ME_{PhonemeType}$	$F(1) = 3.11$
	$p = 0.08$
ME_{List}	$F(3) = 1.18$
	$p = 0.32$
PhonemeType x List	$F(3) = 0.15$
	$p = 0.93$
E7 Words	
M_1 (*SD*)	744.40 (29.03)
M_2 (*SD*)	731.13 (26.05)
M_3 (*SD*)	754.46 (36.26)
M_4 (*SD*)	754.80 (38.88)
L7 Words	
M_1 (*SD*)	762.08 (34.14)
M_2 (*SD*)	745.75 (33.58)
M_3 (*SD*)	761.83 (30.33)
M_4 (*SD*)	740.25 (50.87)
$ME_{PhonemeType}$	$F(1) = 0.84$
	$p = 0.36$
ME_{List}	$F(3) = 1.53$
	$p = 0.21$
PhonemeType x List	$F(3) = 1.13$
	$p = 0.34$

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

Table 32. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), and *p*-value (*p*) of the word frequency data for the words in the four Study 2 stimulus lists.

Statistical Effects	Word Frequency (Kucera-Frances Database)	Word Log Frequency (HAL Database)
E7 Words		
M_1 (*SD*)	70.57 (86.99)	8.43 (2.57)
M_2 (*SD*)	69.00 (72.32)	9.49 (1.75)
M_3 (*SD*)	70.58 (100.75)	8.83 (2.19)
M_4 (*SD*)	191.00 (424.00)	9.93 (2.03)
L7 Words		
M_1 (*SD*)	504.82 (1539.83)	8.65 (3.29)
M_2 (*SD*)	205.27 (460.15)	9.89 (1.92)
M_3 (*SD*)	120.91 (252.32)	8.33 (2.42)
M_4 (*SD*)	63.90 (93.08)	8.20 (3.12)
$ME_{PhonemeType}$	$F(1) = 1.10$	$F(1) = 0.75$
	$p = 0.30$	$p = 0.39$
ME_{List}	$F(3) = 0.54$	$F(3) = 1.30$
	$p = 0.66$	$p = 0.28$
PhonemeType x List	$F(3) = 1.02$	$F(3) = 1.05$
	$p = 0.39$	$p = 0.37$

Note. A 2 x 4 (phoneme type x list) ANOVA was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; 1 – 4 = stimulus lists 1 – 4.

APPENDIX E

ADDITIONAL RESULTS

Table 33. Mean (*M*), standard deviation (*SD*), paired *t*-test (*t*), *p*-value (*p*), and Cohen's *d* (*d*) of accuracy and RT for the strict scoring procedure in the nonword reading task (Study 1).

Dependent Measure	E7 *M* (*SD*)	L7 *M* (*SD*)	Test Statistics
Nonwords (Accuracy)	79.00 (13.98)	66.33 (14.68)	$t(29) = 6.03$ $p < 0.001$ $d = 88.82$
Nonwords (RT in ms)	734.45 (195.26)	735.37 (171.86)	$t(29) = 0.07$ $p = 0.95$ $d = 0.01$

Note. RT is reported for correct responses only. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 34. Mean (*M*), standard deviation (*SD*), paired *t*-test (*t*), *p*-value (*p*), and Cohen's *d* (*d*) of accuracy, RT, and Adjusted RT for the word items in the auditory lexical decision task (Study 1).

Dependent Measure	E7 *M* (*SD*)	L7 *M* (*SD*)	Test Statistics
Words (Accuracy)	89.78 (8.35)	78.33 (13.42)	$t(29) = 5.21$ $p < 0.001$ $d = 107.90$
Words (RT in ms)	1049.00 (127.48)	1133.47 (156.03)	$t(29) = 4.71$ $p < 0.001$ $d = 0.60$
Words (Adjusted RT in ms)	335.77 (124.23)	392.29 (155.89)	$t(29) = 3.18$ $p = 0.003$ $d = 0.41$

Note. RT is reported for correct responses only. E7 = Early-7 stimuli; L7 = Late-7 stimuli.

Table 35. Mean (*M*), standard deviation (*SD*), ANOVA *F* statistic (*F*), *p*-value (*p*), and partial eta squared (η^2) of accuracy, RT, and Adjusted RT for the word items in the Study 2 lexical decision tasks.

Test Statistics	Word Accuracy	Word RT (ms)	Word Adjusted RT (ms)
M_{E7} (*SD*)	87.08 (10.52)	883.95 (267.95)	510.40 (180.08)
M_{L7} (*SD*)	86.15 (10.35)	896.27 (266.55)	519.51 (200.74)
$ME_{PhonemeType}$	$F(1,19) = 0.52$ $p = 0.48$ partial $\eta^2 = 0.03$	$F(1,19) = 2.14$ $p = 0.16$ partial $\eta^2 = 0.10$	$F(1,19) = 1.35$ $p = 0.26$ partial $\eta^2 = 0.07$
M_{Aud} (*SD*)	84.13 (10.55)	1122.19 (131.07)	375.61 (129.43)
M_{Vis} (*SD*)	89.10 (9.71)	658.03 (129.07)	654.30 (129.18)
$ME_{PresentationModality}$	$F(1,19) = 10.74$ $p = 0.004$ partial $\eta^2 = 0.36$	$F(1,19) = 430.84$ $p < 0.001$ partial $\eta^2 = 0.96$	$F(1,19) = 154.88$ $p < 0.001$ partial $\eta^2 = 0.89$
M_{noCA} (*SD*)	88.88 (8.93)	886.01 (261.52)	512.80 (195.65)
M_{CA} (*SD*)	84.35 (11.32)	894.21 (272.94)	517.11 (185.68)
ME_{CA}	$F(1,19) = 10.98$ $p = 0.004$ partial $\eta^2 = 0.37$	$F(1,19) = 0.25$ $p = 0.62$ partial $\eta^2 = 0.01$	$F(1,19) = 0.06$ $p = 0.80$ partial $\eta^2 = 0.003$
Interactions	No Interactions $p \geq 0.38$	No Interactions $p \geq 0.21$	No Interactions $p \geq 0.16$

Note. RT is reported for correct responses only. A 2 x 2 x 2 within-subjects ANOVA design comparing phoneme type x presentation modality x presence of concurrent articulation was used. ME = main effect; E7 = Early-7 stimuli; L7 = Late-7 stimuli; Aud = Auditory presentation; Vis = Visual presentation; noCA = without concurrent articulation; CA = with concurrent articulation.

BIBLIOGRAPHY

Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: silent reading activates inner speech. *Perception & Psychophysics, 59*(7), 1059-1068.

Anthony, J. L., Aghara, R. G., Solari, E. J., Dunkelberger, M. J., Williams, J. M., & Liang, L. (2011). Quantifying phonological representation abilities in Spanish-speaking preschool children. *Applied Psycholinguistics, 32*, 19-49.

Archibald, L. M. D., & Gathercole, S. E. (2005). What's so special about nonword repetition in specific language impairment? *Manuscript submitted for publication*.

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition: a comparison of tests. *Journal of Speech, Language, and Hearing Research, 49*(5), 970-983.

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: more than a phonological short-term memory deficit. *Psychonomic Bulletin & Review, 14*(5), 919-924.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*(3), 189-208.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8*, 485-493.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283-316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445-459.

Berent, I. (1997). Phonological priming in the lexical decision task: regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception and Performance, 23*(6), 1727-1742.

Berent, I., Bouissa, R., & Tuller, B. (2001). The effect of shared structure and content on reading nonwords: evidence for a CV skeleton. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(4), 1042-1057.

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37*(4), 391-403.

Blumberger, J., Sullivan, S. J., & Clement, N. (1995). Diadochokinetic rate in persons with traumatic brain injury. *Brain Injury, 9*(8), 797-804.

Bowey, J. A. (2001). Nonword reptition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*, 441-469.

Bowey, J. A. (2006). Clarifying the phonological processing account of nonword repetition. *Applied Psycholinguistics, 27*, 548-552.

Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 329-340.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, & Hearing Research, 48*(6), 1378-1396.

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes, 16*(2/3), 113-142.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1-40.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences* (revised ed.). New York: Academic Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321.

Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech and Hearing Research, 36*, 1051-1054.

Dollaghan, C., Biber, M., & Campbell, T. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics*, *16*, 211-222.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136-1146.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*(2), 421-436.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19*, 279-309.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading & Writing, 8*, 453-485.

Elbro, C., Borstrom, I., & Petersen, D. K. (1998). Predicting dyselxia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, *33*, 36-60.

Elbro, C., & Pallesen, B. R. (2002). The quality of phonological representations and phonological awareness: A causal link? In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors to functional literacy* (pp. 17-32). Amsterdam: John Benjamins.

Ellis, A. W. (1980). Errors in speech and short-term memory: the effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior, 19*, 624-634.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, & Hearing Research, 43,* 865-878.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition, 23*(1), 83-94.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory*. Hove: Psychology Press.

Goffman, L. (2004). Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, & Hearing Research, 47*(5), 1088-1102.

Goldinger, S. D. (1996). Auditory Lexical Decision. *Language and Cognitive Processes, 11*(6), 559-567.

Graf-Estes, K. G., Evans, J., & Else-Quest, N. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*(1), 177-195.

Guediche, S., Chein, J., & Fiez, J. A. (In review). Similarity effects and working memory: Evidence for a revised model of verbal maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Gupta, P. (2005). Primacy and recency in nonword repetition. *Memory, 13*(3/4), 318-324.

Gupta, P. (2006). Nonword repetition, phonological storage, and multiple determination. *Applied Psycholinguistics, 27*(4), 564-568.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language, 53*, 141-162.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: computational and neural bases. *Brain and Language, 59*(2), 267-333.

Hester, E., & Hodson, B. W. (2004). The role of phonological representation in decoding skills of young readers. *Child Language Teaching and Therapy, 20*(2), 115-133.

Hintzman, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 6*, 312-316.

Hoffmann, K. A. (1982). *Speech sound acquisition and natural process occurrence in the continuous speech of three-to-six year old children.* Unpublished master's thesis, University of Wisconsin-Madison.

Hoffmann, K. A., & Shriberg, L. D. (1982). *Percentage of consonants correct (PCC): A normative study*. Paper presented at the Annual Convention of the American Speech-Language-hearing Assocation, Toronto, Canada.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*(6), 687.

Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment: a deficit in grammar or processing? *Trends in Cognitive Science, 2*(7), 240-247.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 656-674.

Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: accuracy and error data from two nonword repetition tests. *Journal of Speech, Language, and Hearing Research, 53*(6), 1642-1655.

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13*(7/8), 846-863.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17-64). Timonium, MD: York Press.

Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., et al. (2005). Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. *Neuropsychologia, 43*(14), 2068-2083.

Kent, R. D. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 65-90). Timonium, MD: York Press.

Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders, 52*(4), 367-387.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-38; discussion 38-75.

Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*, 123-132.

Mader, J. B. (1954). The relative frequency of occurrence of English consonant sounds in words in the speech of children in grades one, two, and three. *Speech Monographs, 21*, 294-300.

Metsala, J. L., & Chisholm, G. M. (2010). The influence of lexical status and neighborhood density on children's nonword repetition. *Applied Psycholinguistics, 31*, 489-506.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*(2), 338-352.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech, 21*(3), 221-241.

Moore, M. W., Tompkins, C. A., & Dollaghan, C. A. (2010). Manipulating articulatory demands in nonword repetition: A "Late-8" nonword repetition task. *Clinical Linguistics and Phonetics, 24*(12), 997-1008.

Munson, B., Edwards, J., & Beckman, M. E. (2005). 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.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*(5), 1033-1047.

Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience, 30*(25), 8435-8444.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J. L., Nazarian, B., et al. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage, 50*(2), 626-638.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Studies in written language and literacy: Precursors of functional literacy*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Pittsburgh English (2012). Retrieved March 29, 2012, from http://en.wikipedia.org/wiki/Pittsburgh_English

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Redford, M. A., & Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America, 106*(3 Pt 1), 1555-1565.

Rogers, M. A., & Storkel, H. L. (1998). Reprogramming phonologically similar utterances: the role of phonetic features in pre-motor encoding. *Journal of Speech, Language, & Hearing Research, 41*(2), 258-274.

Sander, E. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders, 37*, 55-63.

Schneider, J. (2002). Block letter eye chart, from http://www.i-see.org/block_letter_eye_chart.pdf

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime reference guide. Pittsburgh: Psychology Software Tools Inc.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol Rev, 96*(4), 523-568.

Shriberg, L. D. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Hillsdale, NJ: Lawrence Erlbaum.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36*, 105-140.

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

Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B. (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51*(2), 140-161.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12*, 369-373.

Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London: Whurr.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137-149.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*, 41-78.

Stokes, S. F., & Surendran, D. (2005). Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech, Language, and Hearing Research, 48*(3), 577-591.

Sutherland, D., & Gillon, G. T. (2005). Assessment of phonological representations in children with speech impairment. *Language, Speech, and Hearing Services in the Schools, 36*(4), 294-307.

Sutherland, D., & Gillon, G. T. (2007). Development of phonological representations and phonological awareness in children with speech impairment. *International Journal of Language and Communication Disorders, 42*(2), 229-250.

Taft, M. (2006). Orthographically influenced abstract phonological representation: evidence from non-rhotic speakers. *Journal of Psycholinguistic Research, 35*(1), 67-78.

Tiffany, W. R. (1980). The effects of syllable structure on diadochokinetic and reading rates. *Journal of Speech and Hearing Research, 23*(4), 894-908.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: implications for reading development and theory. *Memory & Cognition, 18*(6), 559-567.

Treiman, R., & Zukowski, A. (1988). Units in reading and spelling. *Journal of Memory and Language, 27*, 466-477.

Vaden, K. I., Halpen, H. R., & Hickok, G. S. (2009). Irvine Phonetic Online Dictionary, Version 2.0, from www.iphod.com

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers, 36*(3), 481-487.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192-212.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive Test of Phonological Processing. Austin, Texas: PRO-ED.

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286-350.

Wolf, M., Vellutino, F., & Berko Gleason, J. (1998). A psycholinguistic account of reading. In J. Berko Gleason & N. Bernstein Ratner (Eds.), *Psycholinguistics* (2nd ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Yates, M. (2009). Phonological neighbourhood spread facilitates lexical decisions. *The Quarterly Journal of Experimental Psychology (Hove), 62*(7), 1304-1314.

Ziegler, J., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*(4), 683-689.