Applied Psycholinguistics **39** (2018), 933–959
doi:10.1017/S014271641800005X

Consonant age of acquisition effects are robust in children's nonword repetition performance

MICHELLE W. MOORE
West Virginia University

Received: January 26, 2017

ADDRESS FOR CORRESPONDENCE
Michelle W. Moore, West Virginia University, P.O. Box 6122, 807B Allen Hall, Morgantown, WV 26506. E-mail: mimoore@mail.wvu.edu

ABSTRACT

The underlying processes of nonword repetition (NWR) have been studied extensively in both typical and atypical development. Most of the research examining long-term memory effects on NWR has focused on lexical and sublexical variables that can only be computed relative to the lexicon of a specific language (e.g., phonotactic probability). Sublexical variables that can be defined without reference to the lexicon (e.g., consonant age of acquisition; CAoA) have received little attention, although recent work has shown a CAoA effect on NWR in young adults by measuring performance differences when the stimuli comprise consonants acquired later versus earlier in speech development. The purpose of this study was to identify whether this sublexical effect occurs earlier in development. Thirty-one typically developing first and second graders completed NWR, nonword reading, and auditory lexical decision tasks. Nonword accuracy and word–nonword discriminability were consistently lower for items comprising later versus earlier acquired phonemes, even after controlling for vocabulary knowledge, but there were no differences in speed measures. Patterns of performance were similar to the CAoA effects observed in young adults from previous work. Results indicate that the sensitivity of NWR performance to these sublexical long-term memory effects occurs in childhood and reflects adultlike patterns of performance.

Keywords: consonant age of acquisition; nonword repetition; phonological memory

Nonword repetition (NWR), which requires the immediate repetition of a spoken nonword, is sensitive to both language experience and language ability as observed across a broad range of groups including adults and children, first and second language learners, and individuals with and without a variety of communication disorders. Because of its clinical utility, significant work has focused on understanding the cognitive linguistic mechanisms engaged during the task in order to better interpret typical and atypical processes involved in language learning. Although NWR was originally viewed as a measure of phonological short-term memory, more current empirical findings and theoretical refinements suggest NWR is a multidimensional measure affected by phonological short-term memory ability, the storage and retrieval of linguistic information from long-term memory

(Archibald & Gathercole, 2006; Gupta, 2006; Gupta & Tisdale, 2009; Rispens & Baker, 2012), as well as other factors. Nearly all of the research examining long-term memory effects on NWR has focused on lexical and sublexical variables that can only be computed relative to the lexicon of a specific language (e.g., phonotactic probability). However, recent work with college students has tested the effects of a sublexical variable (i.e., consonant age of acquisition; CAoA) as a way to understand the long-term phonological knowledge influences in NWR performance that can be defined without reference to the lexicon. To identify whether the effects of this sublexical variable are present in children, the current work extends a previous line of study by testing whether children's NWR, nonword reading, and auditory lexical decision task performances differ when the items comprise consonants acquired later versus earlier in speech development.

Current work utilizing descriptive designs (e.g., Rispens & Baker, 2012) and computational modeling (e.g., Gupta & Tisdale, 2009) has provided strong evidence that NWR performance is affected by factors associated with both phonological short-term memory and long-term phonological knowledge (e.g., Gupta, 2006; Gupta & Tisdale, 2009; Rispens & Baker, 2012). For example, Gupta and colleagues (Gupta, 2003; Gupta & MacWhinney, 1997; Gupta & Tisdale, 2009) proposed a unitary framework for long-term and short-term phonological information to account computationally for the observed relationships between NWR performance, vocabulary acquisition, and immediate serial recall (i.e., a classic verbal working memory task in which participants recall a list of items in their presentation order). Their interactive model has both feedforward and feedback processes between lexical, syllabic, and phonological/phonetic information, such that phonological short-term memory occurs via connections that are weighted based on the network's experience. In other words, phonological short-term memory is functionally represented in the model, but it is influenced by, and not structurally independent from, long-term knowledge. Specifically, Gupta and Tisdale (2009) simulated many of the outcomes observed in behavioral studies reflecting causal effects of short-term and long-term phonological memory in nonword repetition. Thus, it is critical that researchers understand the influences of long-term memory on NWR in order to understand the mechanisms that underlie the association between NWR and language learning (see Moore, Fiez, & Tompkins, 2017, for further discussion).

Research examining the influences of long-term phonological knowledge on NWR has focused primarily on lexical and sublexical effects that are strongly associated with vocabulary knowledge. For example, Gathercole, Willis, Emslie, and Baddeley (1992) reported findings from a longitudinal study examining the relationship between NWR and vocabulary development in children ages 4 through 8. This study found a significant relationship between NWR and receptive vocabulary for children 4, 5, and 6 years of age, but no statistically significant relationship was observed at age 8 after controlling for confounding factors. The direction of the relationship changed with age (i.e., NWR scores at age 4 predicted vocabulary scores at age 5). However, from age 5 to 6 and age 6 to 8, vocabulary scores predicted NWR scores (see Coady & Evans, 2008; Gathercole, 2006, for a review of NWR and vocabulary development in children). Other examples of lexical and sublexical long-term memory effects on NWR include work demonstrating that

NWR performance is improved when the nonwords are more wordlike (Gathercole, 1995; Graf-Estes, Evans, & Else-Quest, 2007), when the stressed syllables of the nonwords are real words (Dollaghan, Biber, & Campbell, 1993, 1995), and when longer nonwords came from more dense phonological neighborhoods (Metsala & Chisholm, 2010).

Previous work also has shown sublexical effects of phonotactic probability on NWR performance, such that NWR performance improves when the nonwords have a higher phonotactic probability (Coady, Evans, & Kluender, 2010; Edwards, Beckman, & Munson, 2004; Munson, 2001; Munson, Kurtz, & Windsor, 2005). The effect from this type of manipulation seems to decrease with age or experience (i.e., vocabulary knowledge; e.g., Edwards et al., 2004; Munson, Edwards, & Beckman, 2005, 2012; Munson, Kurtz, et al., 2005; Roodenrys & Hinton, 2002). For example, Munson et al. (2012) reported that children's repetition performance with non-English biphone sequences (e.g., /fk/) was less accurate compared to repeating sequences that do occur in English (e.g., /ft/, as in *after* or *fifty*). This frequency effect of phonotactic probability decreased with age and was predicted by vocabulary size; however, the effect persisted when speech production accuracy and speech perception were controlled. The authors state two important points relevant to this current discussion: "these findings suggest that the emergence of abstract phonological representations in childhood is yoked to developmental changes in vocabulary size" and "changes across development in the magnitude of the frequency effect appear to be distinct from developmental changes in parametric phonetic knowledge" (p. 298). The first statement illustrates well the hypothesis seen in the literature that latent variables of long-term phonological knowledge are strongly associated with development of the lexicon (see also Metsala & Chisholm, 2010). The authors' second statement suggests that there is speech production/perception knowledge that is distinct from a measure like phonotactic probability. This leaves open the possibility that there are latent variables of long-term phonological knowledge that function independently from vocabulary knowledge.

Until recently, there has been a paucity of work focused on directly manipulating sublexical latent variables of long-term phonological knowledge that could be independent from vocabulary knowledge. A small body of work has explored such sublexical influences in NWR and yielded positive results, thus warranting further consideration. Specifically, sublexical effects in NWR performance have been found using a CAoA manipulation. Moore et al. (2017) state that this manipulation takes advantage of the fact that children master the production of different phonemes at different ages. Consonant acquisition can range from age 3 to age 9 (using a 90% level of acquisition criterion). Phonemes such as /m, n, p/ are typically acquired by age 3 and phonemes such as /s, z, r/ are generally acquired by 7 or later (Smit, Hand, Freilinger, Bernthal, & Bird, 1990).

Unlike phonotactic probability, which by definition is directly associated with the frequency of phoneme occurrence within a language, the relationship between frequency of occurrence and CAoA is not straightforward. Cross-linguistic data suggest that frequency of occurrence plays a role in the order of consonant acquisition, and the degree to which frequency of occurrence contributes to acquisition may depend on factors such as size of vowel inventory, functional load (an effect of minimal contrasts), as well as other factors (Beckman & Edwards, 2010; Stokes &

Surendran, 2005). Thus, frequency may contribute somewhat to CAoA in English; however, it is clear that several later-developing phonemes (i.e., /s, r, l/) are some of the most frequently occurring consonants in English (Mader, 1954; Mines, Hanson, & Shoup, 1978), so frequency of occurrence cannot fully explain the order of acquisition for English consonants. CAoA typically has been associated with the articulatory complexity (i.e., motor demands) of speech sound production (Kent, 1992; Stokes & Surendran, 2005). In Stokes and Surendran's (2005) model of English consonant mastery in which functional load, frequency, and articulatory complexity were considered as possible predictors, articulatory complexity accounted for 40% of the variance (frequency of occurrence and functional load did not significantly contribute to unique variance: 3% and 0%, respectively).

Although CAoA is typically associated with the articulatory complexity of speech sound production (Kent, 1992; Stokes & Surendran, 2005), current leading models of speech production are convergent in depicting articulatory gestural information at an abstract representational level (Dell, 1986; Guenther, Ghosh, & Tourville, 2006; Levelt, Roelofs, & Meyer, 1999). By extension, consonants that are more complex to articulate (i.e., later-developing phonemes) should involve the storage and encoding of a more complex articulatory gesture. One example is Dell's (1986) spread activation model of speech production in which phonetic featural information (i.e., place of articulation, manner of articulation, and voicing) is available during phonological encoding. Dell's model has separate phoneme and featural units, but these individual units are dynamic within the network, with activation occurring in both directions. Because of the interaction between phonemes and featural information, from this model one can postulate that there could be performance differences between early- and late-developing phonemes in part because of the articulatory information available at the level of phonological encoding. Thus, the retrieval of this sublexical information from long-term memory should be less efficient (e.g., slower or less accurate) for later-developing phonemes in various lexical access tasks, even in tasks that do not require articulation.

In studies conducted with normal college students, Moore and colleagues (Moore et al., 2017; Moore, Tompkins, & Dollaghan, 2010) found a robust CAoA effect in NWR. The purpose of Moore et al.'s (2017) work was to determine whether the CAoA effect reflected an articulatory influence on NWR (i.e., an influence of motoric demands) or an influence from a level of phonological representation. If the CAoA manipulation reflected a long-term memory influence on phonological representation and activation within short-term memory, then the CAoA effects would be observed across tasks that vary in perceptual input (e.g., auditory vs. visual input), articulatory demands (e.g., spoken vs. recognition output), and need to maintain phonological information in short-term memory.

In their first of two experiments, Moore et al. (2017) administered three lexical access tasks (NWR, auditory lexical decision, and nonword reading) in which half the stimuli comprised only early-developing consonants and half comprised only late-developing consonants. The stimulus lists were carefully balanced on many lexical and sublexical factors that are often linked with vocabulary acquisition (e.g., wordlikeness, phonotactic probability, etc.) in order to strengthen the claim that any observed CAoA effects are relatively independent from other aspects of word knowledge. Auditory lexical decision and nonword reading tasks were selected to rule out potential confounds due to the CAoA effect (Moore et al., 2010). Auditory lexical decision was selected because it eliminates overt articulatory demands; nonword reading was selected because it eliminates auditory perceptual demands. For Experiment 2, the influence of CAoA effects in lexical access tasks was investigated when articulatory rehearsal was suppressed using the early-late phoneme contrast in auditory and visual lexical decision, with and without concurrent articulation. Across all of the tasks in Experiments 1 and 2, participants on average performed less accurately on nonword items comprising later-developing consonants than on those comprising early-developing consonants. The CAoA effect persisted in lexical decision tasks even when concurrent articulation was used to minimize both overt and covert articulatory processes. The CAoA effect also persisted in nonword reading and visual lexical decision tasks that used visually presented stimuli to minimize auditory perceptual and phonological short-term memory demands. Because the CAoA effect persisted across lexical access tasks that minimized articulatory, short-term memory, and auditory perceptual demands, Moore et al. concluded that the results favored a long-term phonological knowledge view of the CAoA effect over an articulatory view of the effect.

Moore et al.'s findings favoring a long-term phonological knowledge view of the CAoA effect over an articulatory view are consistent with other findings showing modest or no effect of articulatory and motoric demands in NWR performance. Moore and colleagues described previous work assessing articulatory complexity. For example, studies using consonant class (Edwards & Lahey, 1998) and consonant clusters versus singletons (Archibald & Gathercole, 2006) showed no effect on NWR performance in typically developing children (although this was not always the case for the groups of children with language impairment; Archibald & Gathercole, 2006; Bishop, North, & Donlan, 1996; Briscoe, Bishop, & Norbury, 2001; Coady & Evans, 2008). In addition, Archibald, Joanisse, and Munson (2013) found that motorically constrained conditions of NWR (e.g., administration using a bite block) only modestly affected repetition performance in children. Hence, favoring a long-term phonological knowledge view of the CAoA effect over a purely articulation-based influence of the effect on NWR is compatible with other work that does not show significant effects of articulatory demands in NWR.

In summary, the CAoA manipulation is a novel approach to operationalize sublexical influences from long-term memory in NWR performance that are relatively independent from vocabulary knowledge, but thus far the effect has only been examined in adults. An open question is whether the effect is similar in children. Answering this question could provide critical information regarding the nature of the processes utilized in NWR and language learning. If a CAoA effect is present in children, this suggests there may be an important role for sublexical information in language learning that is relatively independent from vocabulary knowledge, and this role is developed early on and continued into adulthood. Alternatively, if a CAoA effect is not present in children, this could suggest that the level of information involved in lexical access tasks may change over the course of development. The study is designed to test for a differential effect of CAoA in typically developing children using nonword repetition and other linguistic tasks.

Specifically, this study will address the following research question: does consonant age of acquisition predict children's performance across a battery of lexical access tasks that vary in their auditory perceptual, articulatory, and short-term memory demands? If the CAoA effect is present, a secondary question will be considered: does consonant age of acquisition predict children's performance across a battery of lexical access tasks after controlling for children's receptive vocabulary ability?

In hypothesizing the outcome for the first question, the robust nature of the CAoA effect should be considered. A significant CAoA effect has been observed with several different samples of young adult participants across a range of linguistic tasks. Although children show more variability than adults in their speech productions when using fine-tuned measures (e.g., kinematic, durational, and spectral measures), they achieve adultlike levels of speech production accuracy by the age of 6 (see Munson et al., 2012, for a review). Thus, even though acoustic and articulatory domains of phonological representation may continue to be fine-tuned over time, it is hypothesized here that long-term phonological knowledge will be sufficiently developed to elicit CAoA effects in children who are 6 and older.

The discussion so far has suggested that the CAoA effect could represent aspects of long-term phonological knowledge that are relatively independent from vocabulary knowledge. One plausible explanation for this that was previously stated is that age of acquisition reflects articulatory gestural information that is distinct from other domains at the representational level. Hence, with the second research question, it is hypothesized that the CAoA effect will persist in performance on the lexical access tasks even after controlling for vocabulary knowledge. Further, if the effect is relatively independent of word knowledge, then the magnitude of the effect should not be different between the children in this study and the young adults tested in the previous work by Moore et al., who presumably would have extensively more lexical knowledge than children.

METHOD

Participants

Participants were recruited from flyers posted around the West Virginia University campus, word of mouth, and in-person invitations offered at various after-school programs in the Monongalia school district (the school district in which West Virginia University is located). Thirty-four participants were initially enrolled in the study; however, 3 were excluded for the following reasons: 1 child did not pass the hearing screening, 1 child did not meet the speech production criteria, and 1 child refused to complete the session tasks. Therefore, the final sample included 31 children (12 females) who were in first or second grade (13 first graders) whose average age was 7.4 years (range 6.3–8.4 years) at the time of enrollment. Parents reported that all of the children were native English monolinguals and that they did not currently have Individualized Education Plans for special education services in school. All passed a pure-tone hearing and a vision screening, and all demonstrated that the targeted early-developing and later-developing consonants sounds were in their phonetic inventory with at least one correct production of each sound articulated during Session 1 of the study. All parents of the participants signed an informed consent, and all children signed an informed assent using procedures approved by the West Virginia University Institutional Review Board. Participants received an age-appropriate storybook at the end of the study for their time and effort.

Experimental tasks

To address the research questions, CAoA stimuli were used in NWR, auditory lexical decision, and nonword reading tasks. The two tasks in addition to NWR were selected to rule out potential confounds to the CAoA effect (Moore et al., 2010). Auditory lexical decision was selected because it eliminates overt articulatory demands by eliciting button-press responses. Nonword reading was selected because it eliminates auditory perceptual demands since the items are presented visually. If CAoA is an effect of long-term phonological knowledge, it should persist across tasks that vary in their input, output, and memory demands.

Technical specifications

For the experimental tasks, participants were seated centered in front of a laptop monitor and viewed visual stimuli from the center of their visual fields. They listened to auditory stimuli through Sony Dynamic Stereo MDR-V6 headphones, and the presentation volume was held constant across all participants. The visual stimuli were centrally presented in white Arial 30-size font against a black background. The auditory stimuli were digitally recorded samples produced by a trained female speaker of Standard American English. They were recorded using an Audio-Technica ATR 20 microphone and Adobe Audition 1.5 software (44100-Hz sampling rate and 16-bit resolution). Verbal responses were digitally recorded using a Logitech USB noise-canceling desktop microphone (model number 980186–0403) and Adobe Audition 1.5 software (44100-Hz sampling rate and 16-bit resolution). Key press responses were recorded using either an RB-730 or an RB-740 Cedrus serial response (SR) box, with subjects using their left index finger to press the second button and their right index finger to press the sixth button in a row of seven buttons on the SR box.

Experimental stimuli

All stimuli and tasks were taken from Moore et al. (2017). The development of the stimuli and tasks will be restated briefly here to facilitate understanding of the current work.

Early- and late-developing phoneme groups

The stimuli comprised either early-developing or late-developing consonant phonemes. The early and late consonant groups (7 phonemes per group; E7 and L7) were taken from the early, middle, and late consonant groups identified by Shriberg and Kwiatowski (1994). In their work, 72 typically developing children aged 3–6 years correctly produced the Early-8, Middle-8, and Late-8 consonants with an average accuracy of 98%, 93%, and 42%, respectively (p. 1108). The "soft g" sound (/ʒ/, as in "beige" and "measure") in Shriberg and Kwiatowski's Late-8 group is difficult to represent in orthographic form and does not occur in the initial position of English words; therefore, it was excluded from the set of late-developing phonemes used in this study. Consonants for the early group were selected from Shriberg and Kwiatowski's Early-8 and Middle-8 groups in order for the E7 group to be more closely matched in articulatory feature distribution to the L7 group, since featural differences were reported as a potential confound in previous work (Moore et al., 2010). The end result was an E7 group comprising phonemes /m, n, p, d, t, f, v/ and an L7 group comprising phonemes /s, z, l, r, ʃ, θ, ð/. The E7 and L7 phonemes were used to construct consonant–vowel (CV) and CVC syllables in the three experimental tasks.

Nonword repetition stimuli

A total of 32 nonwords ranging from one to four syllables in length were used in the task (see Appendix A). A CV structure was used for all nonfinal syllables and a CVC structure was used for final syllables. For half of the stimuli at each syllable length, the consonant phonemes were all E7 phonemes, and for the other half they were all L7 phonemes. Moore et al. (2010, 2017) employed strict criteria to construct the stimuli in order to minimize potential confounds between lists. Factors that were considered included the following: duration of recorded stimuli, phoneme recurrence within a nonword and across the task, phonotactic and biphone probability (Vitevitch & Luce, 2004), and the lexicality of constituent nonword syllables. Vitevitch and Luce's (2004) Web-based phonotactic probability calculator (http://www.people.ku.edu/∼mvitevit/PhonoProbHome.html) provides token-based estimates of phonotactic and biphone probability.

Nonword reading stimuli

A total of 30 stimuli were used for a nonword reading task (these nonwords were taken from one of the four Study 2 nonword lists from Moore et al., 2017; see Appendix A). All of the stimuli were one-syllable with a CVC structure. For half of the stimuli, the consonant phonemes were all E7 phonemes, and for the other half they were all L7 phonemes. The two sets of stimuli were balanced on a number of phonological and orthographic factors shown to affect reading performance: phonological neighborhood density and weighted phonological neighborhood density based on word frequency, phonotactic probability in each phoneme position, biphone probability in each position, number of letters, orthographic neighborhood, mean bigram frequency, summed bigram frequency by position, number of orthographic friends, and consistency ratio.

Auditory lexical decision stimuli

A total of 60 one-syllable, CVC stimuli were used in the task (see Appendix A). Half the items were nonwords, and half were words. For the nonwords, half ($N =$

15) were composed of E7 consonant phonemes and half were composed of L7 consonant phonemes. In contrast to the NWR task, in which constituent syllables were primarily nonwords, most of the nonwords for the auditory lexical decision task were composed of constituent CV words. This was done to encourage participants to listen to all three phonemes before making a decision. The phonological factors that were controlled in the nonword reading stimuli were also controlled here, as well as word frequency for the constituent CVs and duration of the recorded stimuli.

For the words, there were 15 items composed of E7 consonant phonemes, and 12 composed of L7 consonant phonemes. Three words were "mixed" (i.e., composed of an L7 phoneme in the initial position of the word and an E7 phoneme in the final position of the word), because there are relatively few English words that both meet the selection constraints of the study and contain only L7 consonant phonemes. Mixed items were not included in the analysis. 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). The E7 and L7 word lists were balanced on the phonological variables mentioned previously as well as word frequency and duration of the recorded stimuli.

Experimental procedure

After consenting to participate, parents were asked to report any personal history or family history of speech, language, or reading disorder. Following enrollment, the children participated in two separate study sessions that were conducted on different days. All study session procedures were administered by trained and supervised student research assistants, and participants completed sessions individually in a classroom at an after-school program location (i.e., an elementary school) or on campus at West Virginia University.

Session 1 included a pure-tone hearing screening, a subset of items from the Goldman Fristoe Test of Articulation—Second Edition (GFTA-2; Goldman & Fristoe, 2000), and the Wechsler Abbreviated Scale for Intelligence—Second Edition (WASI-II; Wechsler, 2011). Twenty-four items from the GFTA-2 were selected because they contain 1 or more of the 14 E7 and L7 consonant phonemes. The subset of GFTA-2 items and the WASI-II were administered according to the standardized procedures described in their respective manuals.

For Session 2, children completed three experimental tasks followed by the Test of Language Development—Primary Fourth Edition (TOLD-P:4; Newcomer & Hammill, 2008). The experimental tasks were administered on a computer using the E-prime computer program (Schneider, Eschman, & Zuccolotto, 2002). Nonword repetition was administered first to all participants to avoid exposure to other nonwords before completing the task because NWR was the primary focus of the study. Nonword reading and auditory lexical decision followed, with the two tasks counterbalanced across successive participants. The TOLD-P:4 was administered according to the standardized procedures described in the manual.

Task administration for the three experimental tasks was identical to the method described in the Experiment 1 tasks of Moore et al. (2017). The methods are briefly described below.

Nonword repetition task administration. For each trial of the nonword repetition task, participants attempted to repeat aloud an auditory presentation of each nonword. Three practice items were administered. Then, all 8 one-syllable nonwords (4 E7 nonwords, 4 L7 nonwords) were administered, followed by all two-syllable nonwords, and so on. Nonwords within each syllable length were presented in random order for each participant. A red fixation cross was displayed prior to and during the presentation of each stimulus. Then, the cross turned green, prompting the participant to provide the spoken response before the onset of the next stimulus 3 s later.

The phoneme-level scoring described by Dollaghan and Campbell (1998) was used in this current study, such that each phoneme was scored as correct or incorrect compared to the target phoneme. Omissions and substitutions were considered incorrect; distortions were considered correct. Partial responses due to interruption were handled according to the procedures described in Moore et al. (2010). In the few instances ($n = 4$ nonwords) in which a child's repetition was interrupted (e.g., by a hiccup or a question), the phonemes that were scoreable were scored for correctness, and those phonemes that were not attempted due to the interruption were not scored. If an item was not attempted at all, it was not scored (2.5% of the nonwords). These nonresponses were treated as missing data rather than being treated as incorrect because it could not be ruled out that other extraneous factors like inattention were the reason for not attempting to repeat the nonword.

Nonword reading task administration. Participants were asked to read aloud nonword stimuli as quickly and accurately as possible. To begin, 3 practice items were administered and then all 30 nonwords were presented one at a time in random order. For each trial, a white fixation cross appeared and remained on the computer screen until the participant pressed a button on the SR box to elicit a nonword. The nonword remained on the screen until the participant responded aloud.

Two measurements were recorded for this task: whole-word accuracy and reading latency. Responses were marked as correct if the participant pronounced the nonword identically to a target pronunciation (any legal pronunciation of the onset consonant and rime unit). Because of the ambiguity in determining the need for voiced or voiceless "th" in the initial position of nonwords, either phoneme was scored as correct when used in the initial position of any "th" nonword. Reading latency was measured as the duration from the appearance of the nonword to the start of phonation of the response using the spectral and waveform views of the recorded responses in Adobe Audition.

Auditory lexical decision task administration. For the auditory lexical decision task, participants pressed one of two keys to indicate whether each item was a word or a nonword. To begin, a white fixation cross appeared and remained on the screen throughout the duration of the task. Then, the first of 4 practice items was presented, followed by all 60 experimental items one at a time in random order. For each trial, the participants had an unlimited amount of time to respond, although they were instructed to respond as quickly and as accurately as possible. Following each keypress response, there was a 1500-ms intertrial interval before the onset of the next trial. Accuracy and reaction time (RT) were recorded. RT was recorded as the duration from the onset of the stimulus item to the onset of a participant response with the duration of the recorded stimuli subtracted out for each trial.

Scoring reliability

Interrater and intrarater reliability measures were obtained for judgments of accuracy in NWR and nonword reading. Student research assistants independently scored tasks for interrater reliability using participants' digital audio files. A subset of six participants (20% of the sample) was randomly selected for each experimental task. Agreement for judgment of correctness was 85% or greater for E7 and L7 stimuli in both NWR and nonword reading. The primary scorer randomly selected two different subsets of six participants' digital audio files to rescore (one subset per experimental task) for intrarater reliability. The second round of scoring was completed 2 or more months after the initial scoring, and the scorer was blinded to the participants' original scores. Agreement for judgment of correctness was 88% or greater for E7 and L7 stimuli in both tasks.

To measure the reliability of the procedure used to obtain reading latencies for the nonword reading task, undergraduate research assistants independently marked the onset of phonation for the correct responses of six randomly selected participants. The average reading latencies of these six participants differed by less than 50 ms between the first and second scoring ($M_1 = 1191.13$ ms, $SD = 303.54$; $M_2 = 1236.55$ ms, $SD = 366.62$).

RESULTS

Standardized assessments

Descriptive results from the WASI-II and TOLD-P:4 are reported in Table 1. Average scaled and composite scores were within the normal range for all components of the WASI-II and TOLD-P:4; however, some individual scores were >1.5 *SD*s below the normative mean on various subtests of the two standardized tests. Four children scored >1.5 *SD*s below the normative mean on one subtest. In all cases, the subtest was the *oral vocabulary* subtest of the TOLD-P:4 in which participants are asked to define specific words. These four participants did not have any composite scores fall below 1.5 *SD*s of the normative mean. Of note, in the entire sample of children in this study, the *oral vocabulary* subtest had the lowest average score across all 10 WASI-II and TOLD-P:4 subtests (see Table 1). The lower overall scores could be attributed to the lengthiness of the subtest. Anecdotally, it was observed that children seemed more fatigued during this task, which may have contributed to decreased performance. Another consideration is that there is no partial credit given to answers that are "on track" with the correct definition, in contrast to the *vocabulary* subtest of the WASI-II (which is nearly identical in task administration) in which children can receive 0, 1, or 2 points for their responses.

Three children scored >1.5 *SD* below the normative mean on 2 to 3 of the 10 WASI-II and TOLD-P:4 subtests, resulting in composite scores falling below 1.5 *SD* as well. One of these three subjects demonstrated poor performance with the perceptual reasoning components of the intellectual screening, scoring >1.5

Table 1. *Descriptive statistics for WASI-II and TOLD-P:4*

Test composite and scaled scores (SS)	*M*	Range
WASI-II		
Perceptual reasoning composite	96	63–124
Block design SS	9.3	3–14
Matrix reasoning SS	9.5	4–17
Verbal comprehension composite	101	69–124
Vocabulary SS	10.2	6–14
Similarities SS	10.3	2–16
Full scale composite	98.5	74–122
TOLD-P:4		
Picture vocabulary SS	12	6–17
Relational vocabulary SS	8.8	3–14
Oral vocabulary SS	8.0	4–12
Syntactic understanding SS	11.7	6–17
Sentence imitation SS	10.8	6–15
Morphological completion SS	11.1	6–16
Spoken language composite	102.5	77–120

Note: The mean scaled score for WASI-II and TOLD-P:4 subtests is 10, with a $SD = 3$. The mean for all composite scores is 100, with a $SD = 15$. WASI = Wechsler Abbreviated Scale for Intelligence—Second Edition; TOLD-P:4 = Test of Language Development—Primary Fourth Edition.

SD below the mean on the *block design* subtest, the *matrix reasoning* subtest, the perceptual reasoning composite, and the full-scale composite of the WASI-II. One subject primarily had difficulty with tasks that asked the subject to describe how two items are similar, scoring >1.5 *SD* below the mean on the *similarities* subtest and verbal comprehension composite of the WASI-II, and the *relational vocabulary* subtest of the TOLD-P:4. The third subject scored >1.5 *SD* below the normative mean on the *similarities* subtest and the verbal comprehension composite of the WASI-II, the *relational vocabulary* and *oral vocabulary* subtests and the spoken language composite of the TOLD-P:4. Because subtest performance resulted in 1 to 2 composite scores that were outside of the normative range for these three subjects, data analyses for the experimental tasks were performed both with and without their data. The pattern of results was similar for both approaches. The results reported below include the data from these participants.

Plan of analysis

For the primary research question, the data were analyzed using multilevel modeling in HLM v. 7.01 (Raudenbush, Bryk, & Congdon, 2013). The repeated measures were nested at Level 1 within participants at Level 2. For the two models that had continuous dependent variables (nonword reading with reading latency as the dependent variable; auditory lexical decision with RT as the dependent variable), the data were modeled using full maximum likelihood estimation. Reading latency

Table 2. *Descriptive statistics for all variables that were included in the five multilevel models*

	NWR		Nonword reading		Auditory lexical decision	
Variable	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)
CAoA	5527	0.50 (0.50)	930	0.50 (0.50)	1767	0.47 (0.50)
Syllable length	5527	2.89 (1.04)	—	—	—	—
Word type	—	—	—	—	1767	0.53 (0.50)
Probability correct	5527	0.80 (0.40)	930	0.68 (0.46)	1767	0.74 (0.44)
Reading latency	—	—	557	1681.03 (1595.02)	—	—
Reaction time	—	—	—	—	1201	874.75 (465.28)

Note: CAoA: Stimuli with early-acquired consonants were assigned a value = 0, stimuli with late-acquired consonants were assigned a value = 1. Word type: Word stimuli were assigned a value = 0, nonword stimuli were assigned a value = 1. CAoA = consonant age of acquisition; NWR = nonword repetition.

and RT were analyzed for correct responses only. Outliers were considered to be any value that was 2.5 times higher or lower than the upper and lower thresholds (respectively) of the inner quartile range, and were excluded from the analyses.

For the three models with binomial outcomes (probability of correct as the dependent variable for NWR, nonword reading, and auditory lexical decision), the data were modeled using hierarchical generalized linear modeling using adaptive Gaussian estimation with 10 quadrature points. The CAoA variable (E7 or L7 stimuli) was added to each model for each observation. For the NWR model, syllable length and the CAoA × Syllable Length interaction were added as well. For the auditory lexical decision model, word type (word or nonword) and the CAoA × Word Type interaction also were added to the model. All random effects were of substantive interest and therefore were retained for all five models, with two exceptions. The random effect for the CAoA × Syllable Length interaction was not included in the NWR accuracy model and the random effect for the CAoA × Word Type interaction was not included in the auditory lexical decision accuracy model because there was not enough variability in these random effects for the models to converge using the adaptive Gaussian estimation. The descriptive statistics for all five multilevel models are reported in Table 2, and correlations between the variables used in each model are available in online-only Supplemental Tables S.1–S.3.

Nonword repetition

Children's repetitions were analyzed for phoneme-by-phoneme accuracy. One participant was not included in the model due to a recording error. Across the 30 participants included in the analysis, data were not available for 4.1% of all phonemes due to recording error or a child not attempting an item. Because the dependent variable is dichotomous (i.e., each phoneme was scored as correct or incorrect), the parameter estimates are interpreted as exponents that are converted to odds ratios. CAoA was negatively related to accuracy after controlling for the number of syllables (see Table 3 and Figure 1a). The accuracy was always lower when the nonword comprised later-acquired phonemes than when the nonword comprised earlier-acquired phonemes. The results showed a statistically significant interaction between CAoA and syllable length such that the magnitude of the CAoA effect generally increased as the number of syllables increased.

Nonword reading

Nonword reading accuracy and latency were modeled. All 31 subjects were included in the model for accuracy. As described in the NWR model, for the dichotomous dependent variable of accuracy (i.e., probability of correct), the parameter estimates are interpreted as exponents that are converted to odds ratios. CAoA was a statistically significant predictor of accuracy (Table 3 and Figure 1b) such that the likelihood of a correct response was lower for nonwords comprising L7 consonants compared to E7 consonants.

Across the 31 subjects, 6.8% of the trials with correct responses were excluded from the reading latency analysis due to vocalization (e.g., interruption, correction, and/or partial or full repetition of the nonword) that occurred between appearance of the nonword and the participant's final response. Thirty-seven trials (5.8% of trials with correct responses) were treated as outliers (as defined in the Plan of Analysis section) and were deleted, for a total of 557 trials included in the reading latency model. There was no statistically significant relationship between reading latency (for correct responses) and the CAoA variable (Table 3).

Auditory lexical decision

Results for lexical decision accuracy and RT were modeled. All 31 subjects were included in the model for accuracy. As previously stated, the probability of the correct dependent variable is dichotomous, thus the parameter estimates are interpreted as exponents that are converted to odds ratios. CAoA was statistically related to accuracy after controlling for whether each stimulus was a word or a nonword, and there was no CAoA × Word Type interaction (see Table 3 and Figure 1c).

Two participants were excluded from the RT analysis; one was excluded due to recording error, and the second was excluded for inconsistently maintaining the appropriate posture and finger position for reliable RT data. Correct responses only were considered for the RT analysis. Thirty-two trials were treated as outliers (as defined in the Plan of Analysis section) and were deleted, for a total of 1,201 trials included in the RT model. Similar to the nonword reading model for reading latency, there were no significant differences in response times between E7 and L7 items after controlling for whether each stimulus was a word or a nonword (see Table 3).

CAoA effects and vocabulary knowledge

A secondary analysis was completed to address whether the CAoA effect predicts performance on the three tasks after controlling for children's vocabulary

Table 3. *Multilevel model results predicting accuracy, reading latency, and RTs for CAoA effects*

	NWR, accuracy	Nonword reading, accuracy	Nonword reading, reading latency	Auditory lexical decision, accuracy	Auditory lexical decision, RT
Fixed effect	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)
For intercept, π_0					
For intercept, β_{00}	3.90 (0.23)***	1.37 (0.32)***	1212.92 (61.79)***	0.80 (0.11)***	783.50 (36.54)***
For CAoA slope, π_1					
Intercept, β_{10}	–1.69 (0.23)***	–0.61 (0.21)**	55.32 (30.57)	–0.78 (0.15)***	64.00 (44.46)
For syllable length slope, π_2					
Intercept, β_{20}	–0.93 (0.08)***	—	—	—	—
For CAoA × Syllable slope, π_3					
Intercept, β_{30}	0.37 (0.09)***	—	—	—	—
For word type slope, π_2					
Intercept, β_{20}	—	—	—	1.60 (0.29)***	146.31 (39.46)***
For CAoA × Word Type slope, π_3					
Intercept, β_{30}	—	—	—	0.16 (0.25)	–78.03 (62.80)
Random effect	Variance components				
Level 1 (repeated measures)					
Temporal variation, e_{ti}	—	—	101899.50	—	162771.02
Level 2 (within participant)					
Intercept, r_{0i}	0.34***	2.00***	103774.69***	0.02	22580.30***
CAoA slope, r_{1i}	0.05	0.17	4598.34	0.02	12489.26
Syllable length slope, r_{2i}	0.00	—	—	0.04	—
CAoA × Syllable slope, r_{3i}	—	—	—	—	—
Word type slope, r_{2i}	—	—	—	1.21***	16292.89*
CAoA × Word Type slope, r_{3i}	—	—	—	—	43682.59*

Note: CAoA = consonant age of acquisition; NWR = nonword repetition; RT = reaction time. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$.

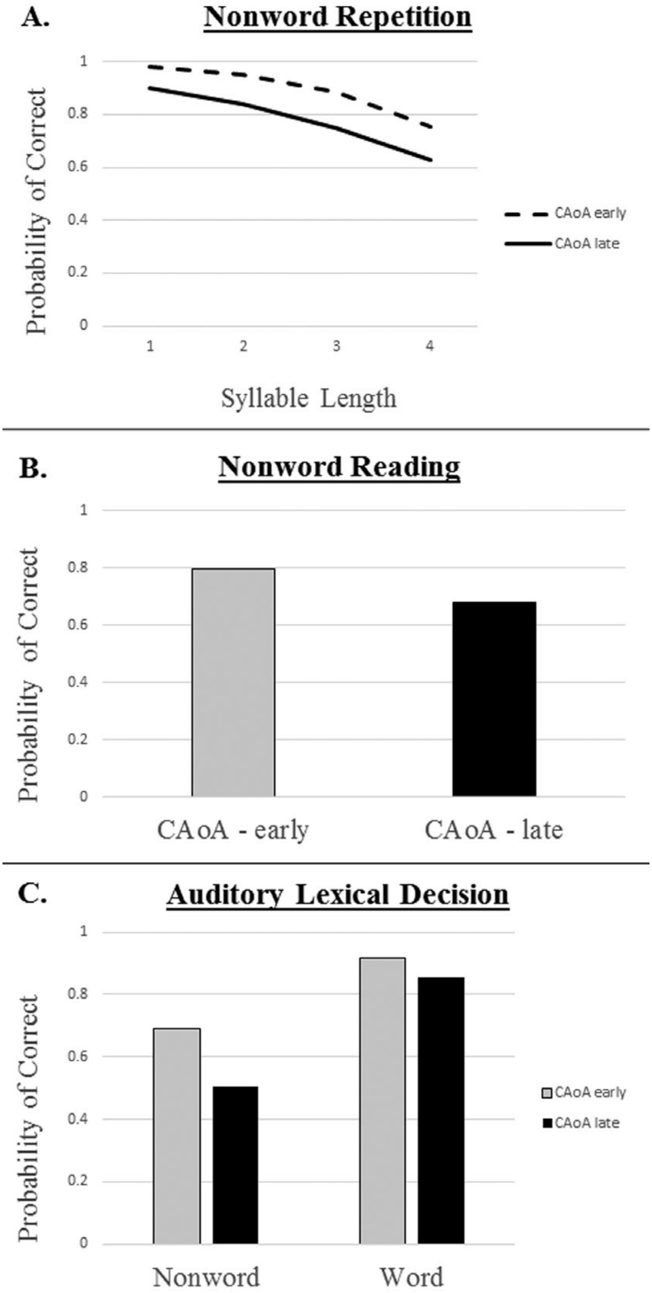

Figure 1. Multilevel models for nonword repetition, nonword reading, and auditory lexical decision. Models suggest robust consonant age of acquisition effects across accuracy measures for all tasks.

knowledge. For the three models showing a significant CAoA effect (i.e., models assessing probability of correct in nonword repetition, nonword reading, and auditory lexical decision), each model was run again adding receptive vocabulary score as a Level 2 factor. All other model parameters were identical to those described above. The scaled scores from the receptive vocabulary subtest of the TOLD-P:4, picture vocabulary, were used in the model. This subtest was used because it is nearly identical in task design as the Peabody Picture Vocabulary Test, a test of receptive vocabulary that has been frequently used in studies assessing the relationship between NWR and vocabulary.[1] All three models had similar outcomes: there was a significant effect of CAoA after accounting for children's receptive vocabulary (NWR CAoA $\beta = -1.68$, $SE = 0.23$, $p = < .001$; nonword reading CAoA $\beta = -0.60$, $SE = 0.20$, $p = .006$; auditory lexical decision CAoA $\beta = -0.78$, $SE = 0.15$, $\mathrm{p} =< .001$). Of note, the CAoA coefficient changed marginally or not at all from the original models in which vocabulary was not included as a Level 2 factor (original models assessing probability of correct: NWR CAoA $\beta = -1.69$, $SE = 0.23$; nonword reading CAoA $\beta = -0.61$, $SE = 0.21$; auditory lexical decision CAoA $\beta = -0.78$, $SE = 0.15$; see Table 3). Thus, the CAoA effect persists to a similar degree of magnitude after controlling for vocabulary knowledge.

If the CAoA effect is relatively independent of word knowledge, it was hypothesized that the magnitude of the effect should not be different between the children in this study and the young adults tested in the previous work by Moore et al. (2017), who presumably would have extensively more lexical knowledge than children. Table 4 presents a comparison of the Cohen *d* effect sizes for each of the measures of the three lexical access tasks used in both studies. Cohen *d* is calculated based on means, standard deviations, and sample sizes, so that comparisons can be made across studies even when different statistical analyses are used.

In both groups, there was a significant CAoA effect on accuracy measures for all three tasks. The magnitude of the effect is generally similar for the two groups on the NWR and nonword reading tasks, even though the two groups had different nonword reading lists. Both groups showed no CAoA effect on speed measures in the nonword reading task and for the nonword items of the lexical decision task. Moore et al. (2017) report a statistically significant CAoA effect for RT on the word items of the auditory lexical decision task, whereas there was no statistically significant effect in the current study for the lexical decision model assessing RT. Cohen *d* values for this word RT measure were small ($d = 0.24$, current study) versus moderate ($d = 0.41$, Moore et al. study). There was one substantial point of difference that should be noted. The auditory lexical decision nonword accuracy effect size was much weaker in the children compared to the adults, despite having similar sample sizes. A possible explanation for this difference is the children's variability in performance on the task. The *SD* for the children's mean performance on auditory lexical decision nonword accuracy for E7 versus L7 items was 14.90% and 14.89%, respectively. In contrast, Moore et al. (2017) report *SD* of 7.20% and 8.60% for mean E7 and L7 nonword accuracy, respectively. Statistically, the increased variance increases the likelihood of making a Type I error and decreases the strength of the Cohen *d* value. Performance variability in lexical access tasks has been observed in younger children previously (Coady & Evans, 2008; Munson et al., 2012), and therefore is not unique to the age of acquisition manipulation.

Table 4. *Comparison of effect sizes for the consonant age of acquisition effect in children (current study) and young adults (Moore et al., 2017) using Cohen d.*

Task	Current study E7 *M* (*SD*)	Current study L7 *M* (*SD*)	Current study Cohen *d*	Moore et al. (2017) Cohen *d*
Nonword repetition				
1-syllable accuracy	95.28 (6.07)	87.22 (8.95)	1.07	0.86
2-syllable accuracy	95.28 (4.55)	83.83 (11.57)	1.32	1.46
3-syllable accuracy	89.19 (10.34)	74.13 (15.02)	1.19	1.09
4-syllable accuracy	72.15 (14.16)	61.38 (15.14)	0.75	1.47
Total accuracy	85.35 (7.56)	73.54 (10.31)	1.33	1.80
			Large effect	Large effect
Nonword reading				
Accuracy	72.47 (24.69)‡	64.52 (23.50)‡	0.34‡	0.45
			Small-moderate effect	Moderate effect
Latency (ms)	1227.67 (379.64)‡	1269.52 (396.80)‡	0.11‡	0.01
			No effect	No effect
Auditory lexical decision				
Word				
Accuracy	68.82 (11.60)‡	50.27 (13.17)‡	1.52‡	1.04
			Large effect	Large effect
RT (ms)	783.73 (195.07)†	839.34 (262.81)†	0.24†	0.41
			Small effect	Moderate effect
Nonword				
Accuracy	88.39 (14.90)‡	83.01 (14.89)‡	0.37‡	1.23
			Small-moderate effect	Large effect
RT (ms)	932.30 (280.60)†	918.44 (289.08)†	0.05†	0.12
			No effect	No effect

Note: $N = 30$ unless otherwise noted. RT = reaction time. †$N = 29$, ‡$N = 31$.

DISCUSSION

Nonword repetition has been studied extensively over the past two decades due in part to its sensitivity in identifying language and reading impairments. It has become so important that poor NWR is viewed as a clinical and phenotypic marker for children with language impairment (e.g., Bishop, 2002; Bishop et al., 1996, 1999; Falcaro et al., 2008; SLI Consortium, 2002, 2004). The clinical utility of NWR has highly motivated the continued research in both typical and atypical development to determine the underlying mechanisms involved in the task and how the task relates to the processes involved with language learning. Although NWR was originally viewed as a measure of phonological short-term memory, the collective body of NWR evidence reveals that repeating a nonword is a multidimensional task requiring, at a minimum, speech perception, phonological encoding and assembly, long-term memory, working memory, and articulation (Coady & Evans, 2008).

The lexical and sublexical influences of long-term phonological knowledge on nonword repetition have been studied extensively, but until recently, there has been little study of sublexical effects that are relatively independent from vocabulary knowledge. Noteworthy findings by Moore et al. (2017) showed an effect of a sublexical variable, consonant age of acquisition, on NWR and other lexical tasks that varied in their perceptual, memory, and output demands in young adults. The purpose of this study was to extend the previous research and to identify whether the effects of this sublexical variable are present earlier in development.

The findings from this current work with first- and second-grade children are similar to the CAoA effect observed by Moore et al. (2017). There was a significant early-late phoneme difference in NWR accuracy, and the magnitude of the difference generally increased as syllable length increased (with the exception of the four-syllable nonwords). The effect size was large for all syllable lengths. The CAoA effect persisted with accuracy measures in the other lexical access tasks in which there was no spoken output (auditory lexical decision) and minimal to no auditory perceptual and phonological short-term memory demands (nonword reading). There was no CAoA effect in the speed measures (i.e., nonword reading latency and lexical decision reaction time).

CAoA as a latent variable of long-term phonological knowledge

One possible interpretation of the observed CAoA effect is that differences in E7 and L7 performance are simply the result of articulatory complexity (i.e., the increased motoric demands required in articulating the later-developing consonants). Although CAoA typically has been associated with articulatory complexity (Kent, 1992; Stokes & Surendran, 2005), the accuracy and speed results seem to work against this interpretation of the CAoA effect. This articulatory view of the results does not explain the observed CAoA effect in auditory lexical decision accuracy in which there is no spoken output. In NWR in which there is a short-term memory component to the task, increased articulatory complexity should decrease the rate of speech production during rehearsal, thus making short-term storage of phonological information more prone to decay. This explanation of the effect in NWR cannot account easily for the observed CAoA effect in accuracy measures for the one-syllable NWR items nor the visually presented, one-syllable nonword reading items. Further, because the basis for this articulatory explanation in the short-term memory capacity of NWR is explained by decreases in speech rate during rehearsal, early-late phoneme differences in speed measures might be expected. Thus, the lack of early-late phoneme differences in reading latency and reaction time in the current study seems to work against the interpretation that the CAoA effect is purely the result of increased motoric demands.

An articulatory view of the CAoA effect in nonword repetition seems to be inconsistent with other work examining the articulatory influences in NWR, as referenced above. That is, several studies assessing articulatory complexity have shown no significant effect of complexity on NWR in typically developing children (Archibald & Gathercole, 2006; Archibald et al., 2013; Edwards & Lahey, 1998), although this may not be the case in children with language impairment (Archibald & Gathercole, 2006; Bishop et al., 1996; Briscoe et al., 2001; Coady & Evans, 2008). Hence, explaining the CAoA effect in NWR as a purely articulation-based influence on performance seems incongruous, both with the results of this current study and with other findings.

An alternative explanation for the CAoA effect is that consonants that are more complex to articulate (i.e., later-developing phonemes) should involve the storage and encoding of a more complex articulatory gesture, which could be less efficient (e.g., slower or less accurate) to access for various linguistic tasks. In typical speech misarticulations, phoneme swaps occur in similar syllable positions of nearby syllables or are caused by feature overlap between a phoneme competitor (Dell, 1986). Accordingly, later-developing phonemes that are processed less efficiently are more vulnerable to a greater number of phoneme misselections due to the decreased ability to resolve competing information. This view of CAoA as a latent variable of long-term phonological knowledge is compelling because it is able to account for the observed CAoA effect in one-syllable stimuli in which there were little to no phonological short-term memory demands, as well as account for observed effects with visual stimuli and tasks with no spoken output.

If the later-developing phonemes have less efficient representations so that it might be more difficult to "rule out" competing information during phoneme selection, it may be surprising that there were no speed differences between E7 and L7 items. One possible explanation for this is that participants adapted to the experimental tasks by putting more emphasis on speed than accuracy. Another possibility is that the brevity of the CVC stimuli in nonword reading and lexical decision account for the lack of speed 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. Future work could examine this idea using longer stimulus items in various speed measures.

CAoA is relatively independent of vocabulary knowledge

In Moore et al.'s (2017) previous work examining the CAoA effect in NWR and other lexical access tasks, their stimulus lists were carefully balanced on many lexical and sublexical factors that are often linked with vocabulary acquisition (e.g., wordlikeness, phonotactic probability, etc.) in order to strengthen the claim that any observed CAoA effects are relatively independent from other aspects of word knowledge. This study took these efforts one step further by examining the impact of lexicon size on the CAoA effect. There were two ways size of the lexicon was addressed: by controlling for children's vocabulary knowledge in the statistical analyses, and by comparing the magnitude of the CAoA effect in children (current study) to the same effect in young adults (Moore et al. study). The results from the statistical models showed that there was no change in the CAoA effect in any of the three tasks when receptive vocabulary was entered as a factor. The CAoA coefficient was nearly identical in the models that included the vocabulary measure compared to the models that did not include the measure. In the second approach for considering lexicon size, Cohen *d* values were computed for all tasks and compared to those from Moore et al. The overall magnitude of the CAoA effect across the two age groups was similar, suggesting that the CAoA effect in children reflects adultlike patterns of performance. To address the few instances of discrepancies between the younger and older cohorts, future work could further evaluate these findings using longitudinal study designs. Taken together, the results are consistent with the proposition that the CAoA effect in NWR may arise from aspects of long-term phonological knowledge that have not been tapped into previously with other sublexical variables and that is relatively independent from vocabulary knowledge.

Other considerations

To better understand the complex relationship between lexical and sublexical levels of representation, future work might consider other phonological variables that may function relatively independently from vocabulary knowledge. Although previous work has observed phonotactic probability effects in NWR that seem to be mediated by vocabulary knowledge (e.g., Munson et al., 2012), one point of consideration is whether a token versus type estimate of phonotactic probability would relate differently to lexicon size. While it might be expected that type frequency effects would be associated with vocabulary knowledge, the nature of the relationship between vocabulary knowledge and token frequency is less clear. Future work could experimentally manipulate type and token estimates of phonotactic probability to test for a differential effect on NWR performance and relationship to lexical variables.

This study focused on the possibility of the CAoA effect operating as a latent variable based on speech production models that depict articulatory gestural information at an abstract representational level. However, there are alternative considerations to explain CAoA as a sublexical measure of long-term phonological knowledge. For example, Moore et al. (2017) suggest an alternative view (which they note as tenuous) in which the several-year span between mastery of early-learned and late-learned phonemes could affect the structure of neural networks and thus the quality and retrieval of phonological representations. For semantic networks, 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 the neural organization of phonological information, then it could explain the varied quality in the speech–language architecture. Perhaps the ideas described here could be examined further by extending previous functional magnetic resonance imaging research that has investigated the neural substrates involved in the speech production network (Peeva et al., 2010) or by utilizing the computational models that have attempted to model the different types of phonological knowledge simultaneously (see Oudeyer, 2005, and Redford & Miikkulainen, 2007, in Munson et al., 2012). For example, Peeva et al. (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 explore the CAoA effect within the Peeva et al. paradigm to determine if the locus and extent of activation varies between phoneme groups.

Future work also could utilize this study's experimental design to adjudicate between the various theoretical accounts of deficits that have been associated with language impairment. For example, a positive effect of sublexical influences from long-term memory that are relatively independent from word knowledge could suggest that language impairment in children may not be strictly from a limited capacity of short-term memory (cf. Gathercole, 2006) or from limited vocabulary knowledge (e.g., Metsala & Chisholm, 2010). Many studies have used descriptive designs to identify the contribution of various speech and language factors involved in the NWR deficits observed in children with language and reading impairment, but there is a paucity of experimental research in this area. Thus, employing the systematic approach used in this study has potential theoretical and clinical implications (Moore et al., 2017).

Conclusion

Children are sensitive to the persistent effect of a sublexical CAoA variable across a variety of linguistic tasks varying in memory, auditory perceptual, and articulatory demands. The results are consistent with the proposition that the CAoA effect in NWR may arise from aspects of long-term phonological knowledge that have not been tapped into previously with other sublexical variables and that is relatively independent from lexical-level information. When the results from this study were compared to similar work conducted with young adults, the observed pattern of results indicated that this long-term phonological knowledge acquired early in development seems to reflect adultlike patterns of performance, at least when processing novel information such as nonwords.

APPENDIX A

Nonword repetition stimuli

Syllable length	E7 stimuli	L7 stimuli
One Syllable	/faɪp/	/zaɪθ/
	/vom/	/θeʃ/
	/taʊd/	/lɔɪs/
	/dɔɪf/	/raʊð/
Two Syllable	/tævɑm/	/sæθɑʃ/
	/maʊnɑv/	/ʃaʊzeð/
	/vefum/	/θulaʊʃ/
	/nɔɪtef/	/ðɔɪzɑl/
Three Syllable	/dævepɔɪn/	/θɑʃɔɪzos/
	/mɔɪpæfaʊn/	/sɑlæður/
	/vutaʊmæf/	/θuræðɔɪʃ/
	/faɪdɑvop/	/ʃæθɔɪraʊs/
Four Syllable	/dɑvupɔɪtæf/	/ðusæzelɔɪθ/
	/nɔɪfaʊmævɑt/	/rælaʊʃɔɪzɑs/
	/tedæfumɔɪn/	/zaɪθɔɪʃæsar/
	/maʊdɔɪnɑpæv/	/ʃaʊθezolæð/

Auditory lexical decision stimuli

Early-7 List		Late-7 List	
Nonwords	Words	Nonwords	Words
/don/	/din/	/laɪʃ/	/liʃ/
/daɪt/	/daɪv/	/lɛʃ/	/lɛs/
/fæp/	/fæn/	/loθ/[1]	/lor/
/fiv/	/fit/	/rus/	/res/
/fɑm/	/fom/	/roʃ/	/roz/
/mɪv/	/mɪt/	/ʃul/	/ʃiθ/
/mup/	/mɑp/	/ʃis/	/ʃol/
/nef/	/nep/	/sɔɪθ/	/sɪl/
/naʊd/	/nod/	/soð/	/suð/
/pɛm/	/pɛt/	/θir/	/θim/*
/paʊf/	/puf/	/θaʊz/	/θaɪz/
/tɔɪm/	/tim/	/ðaɪs/	/ðaɪn/*
/tɑv/	/tɑp/	/ðez/	/ðoz/
/vaɪf/	/vaɪn/	/zɛl/	/zil/
/vid/	/vɔɪd/	/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.

[1] The phoneme sequence /loθ/ is a real word, “loath.” However, it was treated as a nonword due to its extremely rare use in current everyday American English.

Nonword reading stimuli

	Nonwords	
StimulusLists	Spelling	Pronunciation
Early-7 list	dight	/daɪt/
	dute	/dut/
	fape	/fep/
	foon	/fun/
	mide	/maɪd/
	tep	/tɛp/
	veem	/vim/
	naid	/ned/
	noop	/nup/
	pime	/paɪm/

	Nonwords	
StimulusLists	Spelling	Pronunciation
Late-7 list	peaf	/pif/
	taff	/tæf/
	toove	/tuv/
	vade	/ved/
	vome	/vom/
	zool	/zul/
	loath	/loθ/[2]
	luss	/lʌs/
	rall	/rɔl/
	suzz	/sʌz/
	saze	/sez/
	seash	/siʃ/
	shez	/ʃɛz/
	shile	/ʃaɪl/
	thar	/θɑr/
	thil	/θɪl/
	thure	/ður/
	thush	/ðʌʃ/
	zel	/zɛl/
	soth	/sɔθ/

[2] The phoneme sequence /loθ/ is a real word, "loath." However, it was treated as a nonword due to its extremely rare use in current everyday American English.

ACKNOWLEDGMENTS

This study was funded in part by a Grace Clements Research Award for Communication Sciences and Disorders at West Virginia University. A sincere thank you to the research team in the Language and Literacy Lab; to Julia Hamilton, Director of Extended Day Activities; to the site coordinators and volunteers involved with Mountaineer Boys and Girls Club and Afternoon Adventures; and to the participants and their families for all of the various contributions of time and effort related to this project.

NOTE

1. Note that the Peabody Picture Vocabulary Test was not administered in this study to avoid redundancy since there were already five vocabulary subtests as part of the WASI-II and TOLD-P:4, including the *picture vocabulary* subtest of the TOLD-P:4, which is very similar in task design to the Peabody Picture Vocabulary Test. The WASI-II and TOLD-P:4 were selected in order to obtain a broad picture of cognitive linguistic functioning with an IQ score and overall spoken language score.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S014271641800005X

REFERENCES

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

Archibald, L. M. D., Joanisse, M. F., & Munson, B. (2013). Motor control and nonword repetition in specific working memory impairment and SLI. *Topics in Language Disorders, 33*, 255–267.

Beckman, M. E., & Edwards, J. (2010). Generalizing over lexicons to predict consonant mastery. *Laboratory Phonology, 1*. doi:10.1515/labphon.2010.017

Bishop, D. V. M. (2002). The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders, 35*, 311–328.

Bishop, D. V. M., Bishop, S., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research, 42*, 155–168.

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*, 391–403.

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.

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 and Communication Disorders, 43*, 1–40. doi:10.1080/13682820601116485

Coady, J. A., Evans, J. L., & Kluender, K. R. (2010). Role of phonotactic frequency in nonword repetition by children with specific language impairments. *International Journal of Launguage and Communication Disorders, 45*, 494–509.

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

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

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

Dollaghan, C. A., & Campbell, T. F. (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*, 421–436. doi:10.1044/1092-4388(2004/034)

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.

Falcaro, M., Pickles, A., Newbury, D. F., Addis, L., Banfield, E., Fisher, S. E., . . . Conti-Ramsden, G. (2008). Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment. *Genes, Brain and Behavior, 7*, 393–402. doi:10.1111/j.1601-183X.2007.00364.x

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*, 83–94.

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

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887–898. doi:10.1037/0012-1649.28.5.887

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

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*, 177–195.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*, 280–301. doi:10.1016/j.bandl.2005.06.001

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56A*, 1213–1236.

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

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267–333. doi:10.1006/brln.1997.1819

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language, 61*, 481–502.

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.

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

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.

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

Moore, M. W., Fiez, J. A., & Tompkins, C. A. (2017). Consonant age-of-acquisition effects in nonword repetition are not articulatory in nature. *Journal of Speech, Language, and Hearing Research, 60*, 3198–3212. doi:10.1044/2017_JSLHR-L-16-0359

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*, 997–1008.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research, 44*, 778–792.

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*, 61–78.

Munson, B., Edwards, J., & Beckman, M. E. (2012). Phonological representations in language acquisition: Climbing the ladder of abstractions. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.), *Oxford handbook of laboratory phonology* (pp. 288–309). Oxford: Oxford University Press.

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*, 1033–1047.

Newcomer, P. L., & Hammill, D. D. (2008). *Test of Language Development-Primary—Fourth Edition (TOLD-P:4)*. Austin, TX: PRO-ED.

Oudeyer, P.-Y. (2005). How phonological structures can be culturally selected for learnability. *Adaptive Behavior, 13*, 269–281.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J. L., Nazarian, B., & Alario, F. X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage, 50*, 626–638. doi:10.1016/j.neuroimage.2009.12.065

Raudenbush, S. W., Bryk, A., & Congdon, R. (2013). HLM 7.01 for Windows [Hierarchical linear and nonlinear modeling software]. Los Angeles: Multivariate Software.

Redford, M. A., & Miikkulainen, R. (2007). Effects of acquisition rate on emergent structure in phonological development. *Language, 83*, 737–769.

Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research, 55*, 683–694.

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 29–33. doi:10.1037/0278-7393.28.1.29

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

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

SLI Consortium. (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American Journal of Human Genetics, 70*, 384–398.

SLI Consortium. (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American Journal of Human Genetics, 74*, 1225–1238. doi:10.1086/421529

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779–798.

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*, 577–591. doi:10.1044/1092-4388(2005/040)

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*, 481–487.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence—Second Edition (WASI-II)*. San Antonio, TX: Psychological Corporation.