Phonological Encoding in Speech Sound Disorder: Evidence from a Cross-Modal Priming Experiment

Benjamin Munson and **Miriam O.P. Krause**

Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis USA

Abstract

Background—Psycholinguistic models of language production provide a framework for determining the locus of language breakdown that leads to Speech Sound Disorder (SSD) in children.

Aims—This experiment examined whether children with SSD differ from their age-matched peers with typical speech and language development (TD) in the ability to phonologically encode lexical items that have been accessed from memory.

Methods & Procedures—Thirty-six children (18 with TD, 18 with SSD) viewed pictures while listening to interfering words (IW) or a nonlinguistic auditory stimulus presented over headphones either 150 ms prior to, concurrent with, or 150 ms after picture presentation. The phonological similarity of the IW and the pictures' names varied. Picture-naming latency, accuracy, and duration were tallied.

Outcomes & Results—All children named pictures more quickly in the presence of an IW identical to the picture's name than in the other conditions. At the +150 ms stimulus onset asynchrony, pictures were named more quickly when the IW shared phonemes with the picture's name than when they were phonologically unrelated to the picture's name. The size of this effect was similar for children with SSD and children with TD. Variation in the magnitude of inhibition and facilitation on cross-modal priming tasks across children was more strongly affected by the size of the expressive and receptive lexicons than by speech-production accuracy.

Conclusions & Implications—Results suggest that SSD is not associated with reduced phonological encoding ability, at least as it is reflected by cross-modal naming tasks.

Keywords

Speech Sound Disorder; Phonological Encoding; Naming

The task of naming a line drawing of a familiar object seems naïvely to be an effortless one. A closer inspection shows that many different cognitive processes are involved. First, the picture must be perceptually encoded and identified as a known object. Once a picture is identified, the correct name for the picture must be selected from the mental lexicon. After this, the sound structure of the word must be activated. After the sounds that comprise a word have been activated, they are translated into a set of articulatory movements needed to vocalize the picture's name.

The current study examines how three- to seven-year old children associate the semantic-conceptual representation for a picture with the phonemes needed to produce the picture's name, a process called *phonological encoding* in the two-stage model of language production described in Levelt (1989). Specifically, we compare how phonological encoding differs between children with typical speech and language development (TD) to children with Speech Sound Disorder (SSD). Children with SSD produce many speech-sound errors in the absence of a clear medical etiology, such as hearing impairment, neuromuscular disorder, or psychosocial impairment. Children with severe SSD may make numerous errors, are often highly unintelligible to unfamiliar listeners, and usually require speech and language therapy to achieve intelligible speech. One widely cited study of the prevalence of SSD found that it occurs in approximately 3.8% of preschool children, using a conservative criterion for identification (Shriberg, Tomblin, & McSweeny, 1999). Prevalence estimates in other studies vary widely, from as low as 2% to as high as 25% (Law, Boyle, Harris, Harkness, & Nye, 2000).

There is no consensus on the underlying causes of SSD. A large body of research has demonstrated that SSD is associated with deficits in speech perception (e.g., Edwards, Fox, and Rogers, 2002; Rvachew & Jamieson, 1989), speech-motor control (e.g., Edwards, 1992; Towne, 1994), and phonological awareness (Bird, Bishop, & Freeman, 1995; Rvachew & Grawburg, 2006). Moreover, at least two recent studies have examined the ability of children with SSD to complete a task of executive function requiring them to flexibly switch among different rules. Children with certain types of severe SSD were found to perform more poorly than their peers (Dodd & McIntosh, 2010; Crosbie, Holm, & Dodd, 2009). A recent large-scale factor-analytic study of children with SSD examined these children's performance on a large set of clinical measures of speech production, speech processing, and morphosyntax (Lewis et al., 2006). This study found that the tasks grouped into two factors, one on which tasks that manipulated the sound structure of language loaded heavily, and one on which the morphosyntactic tasks loaded heavily. Another exploratory study on SSD found that children clustered into two groups, one of whom appeared to have a motoric impairment and one of whom had a more general delay (Vick, Campbell, Shriberg, Green, Truemper, Rusiewicz, & Moore, 2015).

Studies of long-term outcomes of children with early childhood SSD provide another type of evidence for subgroupings of children with SSD. Lewis, Freebairn, Tag, Ciesla, Iyengar, Stein, and Taylor (2015) showed that long-term outcomes of children with SSD are worse for children with a co-morbid language impairment. Preston, Hull, and Edwards (2013) found that children whose errors were atypical for English learners (such as the backing of alveolar consonants to velars rather than the relatively more common error of fronting velars to alveolars) had poorer phonological awareness and literacy at school age than did children with only typical errors. In contrast, Wren, Roulstone, and Miller (2012) found few differences among groups of school-aged children with persistent speech sound disorders of varying severity.

Relatively less research has examined the specific components of the production process that are deficient in SSD. Consider, for example, the tasks used by Lewis et al. The measures of production in that study were clinical assessments of speech production accuracy (i.e.,

instruments designed largely to describe children's language rather than to determine etiology), or measures of repetition, which test both production and perception simultaneously. The consequence of this is that the precise locus of the breakdown in SSD is still largely unknown. There is a robust literature in psycholinguistics that examines the cognitive processes involved in speech production using experimental manipulations, but this work has not been widely applied in the context of speech disorders. The current study addresses this lack of research by examining how children with SSD differ from children with TD in a picture-naming task using cross-modal primes. To motivate the use of this task, the remainder of the introduction reviews how cross-modal naming tasks have informed our understanding of language production.

The production of a spoken utterance requires multiple cognitive-linguistic and perceptual-motor processes. Schriefers, Meyer, and Levelt (1990) proposed that language production progresses in a series of processes, from the intention to communicate to the articulatory implementation of the message. Their two-stage model of one of these processes, lexical access, begins with a semantic stage, in which lexical items are coded in terms of syntactic and semantic properties. This stage is followed by a retrieval of the phonological structures of words. Schriefers et al. (1990) present evidence for the two-stage model of lexical access from a cross-modal picture-word interference (PWI) task. In a PWI task, people attempt to name pictures while listening to an interfering stimulus that they are instructed to ignore. "Cross-modal" refers to the combination of visual and auditory stimuli in the procedure. In Schriefers et al., Dutch-speaking participants listened to interfering words (henceforth IWs) on headphones, while simultaneously looking at slides projected onto a screen. The IWs are presumed to activate the same set of abstract lexical and phonological representations that the pictures activate. Hence, the IWs' phonological and lexical-semantic similarity to the lexical and semantic characteristics of each picture's name can be varied. The facilitative or inhibitory effects of IWs on picture naming can be examined.

In the PWI task, the timing of the interfering words relative to the picture's appearance can also be varied. IWs are often presented either simultaneously with the onset of the picture (stimulus onset asynchrony [SOA] = 0 ms), before the onset of the picture (in Schriefers et al. and in this article, 150 ms prior, i.e., SOA = −150 ms), or after the onset of the picture (in Schriefers et al. and in this article, 150 ms after, i.e., SOA = +150 ms). The use of different SOAs allows researchers to draw inferences about stages of language formulation. Effects that happen at negative SOAs are presumed to reflect early stages of language production, while those that happen at positive SOAs are presumed to reflect later stages of the production process. Four types of interfering words were presented by Schriefers et al. These were either neutral (i.e., a condition in which an auditory stimulus is presented that is not predicted to affect the naming response systematically, in the study by Schriefers et al. *blanco* [the Dutch word for *blank*]), phonologically related to, semantically related to, or phonologically and semantically unrelated to the picture's name. A fifth condition in which no IW was presented was also included. Mean picture-naming reaction times varied systematically as a function of SOA and IW type. At the −150 ms SOA, the mean picture-naming reaction time was significantly longer when semantically related IWs were presented than when unrelated IWs were presented. No effects of phonologically similar IWs were found at the −150 ms SOA. Schriefers et al. conjectured that interfering words

played in the −150 ms condition would be perceived by participants at about the time when they were completing lexical accessing of picture's name. The fact that picture-naming times were inhibited by semantically related stimuli but not phonologically related stimuli at −150 ms SOA is evidence that there is an early stage of processing in which only semantic information is activated. At the later SOAs, this semantic interference was absent, but there was a pattern of phonological facilitation: mean reaction times were significantly faster when the phonologically related words were presented than when the unrelated words were presented. This was interpreted as evidence that the semantic stage is followed by a phonological stage; that is, that the IWs activated a set of phonological representations. When these were the same as the phonological representations needed to produce the picture's name, picture naming was faster than when the IW did not activate phonological representations needed to produce the picture's name. In the remainder of the paper, we refer to this phonological stage as *phonological encoding*.

Much subsequent research has examined production processes using PWI. A comprehensive review of these studies is outside of the scope of this article, but can be found in Levelt, Roelofs, and Meyer's (1999) detailed review on models of language production and methods used to study stages in language production. It is notable that computational instantiations of this model have been useful in the study of communication impairments. For example, they are able to predict attested patterns of naming errors in individuals with aphasia (Schwartz, Dell, Martin, Gahlo, & Sobel, 2006). In general, studies subsequent to Schriefers et al. have differed in the strength of the evidence that semantic and phonological stages are strictly sequential. For example, Starreveld and La Heij (1996) present evidence from a cross-modal PWI task using written interfering words that suggests overlap between the semantic and phonological stages. However, studies have been relatively consistent in their finding that semantically related IWs interfere with word production; that phonologically related IWs facilitate word production; and that semantic processing occurs before phonological processing. Consequently, the PWI task can be used to assess the integrity of the semantic and phonological stages of picture naming, as well as their time course.

Two studies used the PWI paradigm to examine the phonological stage of typically developing children's picture naming. The first of these is Brooks and MacWhinney's (2000) study of picture-word interference in children and adults. Brooks and MacWhinney used a picture-word interference task very similar to that of Schriefers et al. (1990). Three groups of children participated, aged five, seven, and nine years, along with a group of college-aged adults. Five types of interfering stimuli were used: onset-related, rime-related, identical, and phonologically unrelated IWs, and a neutral condition with the IW *go*. As in earlier studies, three SOAs were used: −150 ms, 0 ms, and +150 ms. As in many studies of children's reaction times (e.g., Kail, 1991), Brooks and MacWhinney found that mean reaction times were faster for adults than for children. Picture naming response latencies were longer for all non-identical interfering words than for the identical interfering words and the neutral *go* stimulus.

The critical finding of Brooks and MacWhinney (2000) was that the influence of phonological similarity between the IW and the picture name on picture-naming reaction times was age-dependent. When the IW shared an onset with the picture's name, a

phonological facilitation effect was found, as shown by the faster reaction times of the onset-related IWs as compared to the phonologically unrelated IWs. This was evident at the 0 ms and +150 ms SOAs for all three groups of children, but only at the 0 ms SOA for adults. This complements Schriefers et al.'s (1990) findings of facilitation for phonologically-related words with adult participants at the 0 ms and +150 ms SOAs, though it should be noted that the phonologically related IWs used by Schriefers et al. did not control for the type of phonological relatedness. A different pattern of results was found when comparing the rime-related IWs to unrelated IWs. Here, only the five- and seven-year-old children showed a facilitative effect of phonological relatedness; no effect was found for older children or for adults.

Brooks and MacWhinney (2000) interpreted their findings to indicate that children's language-formulation processes mature after age seven in a way that improves the efficiency of language production processes. The facilitative effect of rime-related IWs suggests that five- and seven-year-old children do not begin to create production plans for picture names until after they have heard the entire IW. In contrast, the absence of such an effect in nine-year-old children and adults suggests that preparation for picture naming begins before the offset of IW, perhaps even before the offset of the IW's onset consonant. Consequently, IWs that share word-initial input help speed the creation of output plans in nine-year-olds and adults, while those that don't share word-initial information do not speed the formation of these plans. Only five- and seven-year-old children show a facilitative effect of rime-relatedness, because their language-formulation systems are sufficiently slow and inefficient that they can benefit from the additional phonological activation provided by a rime-related IW. Similar findings are reported by Jerger, Martin, and Damian (2002), who conducted a picture-word interference task with children aged five to seven. As in Brooks and MacWhinney (2000), Jerger et al. (2002) found that the naming latency decreased when the interfering stimulus (here, a CV sequence identical to the initial CV in the picture's name) and the picture name shared onsets, while latency increased when onsets were different. As in other work, the relationship between the type of interfering word and naming latencies varied as a function of SOA.

Given that the PWI task is able to examine different stages of language production, one can hypothesize that performance on this task would differ between individuals with and without impairments in language production. One might hypothesize the children with a language impairment would perform like younger, typically developing children, i.e., they might show a larger facilitative effect of rime-related IWs on picture naming than age-matched typically developing children. We might also predict that children with impairments would perform similarly to age-matched typically developing children on naming when IWs share onsets, because both older and younger groups in Brooks and MacWhinney showed a similar degree of facilitation from shared onsets at 0 ms and 150 ms.

Two recent studies used the PWI task to examine naming in children with and without specific language impairment (SLI, i.e., morphosyntactic impairments in the absence of a clear predisposing condition). Seiger-Gardner and Brooks (2008) examined the performance of 8- to 12-year-old children with and without SLI on the experiments described in Brooks and MacWhinney. They found similar patterns of influence for onset-related IWs on naming

latencies for the two groups at late SOAs. However, group differences in a rime-related condition did not follow this prediction: children with typical language development did not show facilitative effects of rime-related IWs at late SOAs, as would be predicted from Brooks and MacWhinney, given their age. However, the children with SLI also did not show these effects. That is, the children with SLI did not behave like younger, typically developing children.

Seiger-Gardner and Schwartz (2008) conducted a PWI experiment with 8- to 12-year-old children with SLI. Their study differed from earlier ones in that they examined the effect of both semantically and phonologically related IWs on naming latency at a wider range of SOAs. The phonologically related IWs shared the onset with the picture's name. Seiger-Gardner and Schwartz included later SOAs than Seiger-Gardner and Brooks, to test the possibility that priming effects linger in children with SLI due to their overall slower speed of processing. Seiger-Gardner and Schwartz found that children with SLI showed patterns of phonological facilitation that were qualitatively similar to those of typically developing children, particularly at a SOA of +150 ms, i.e., the SOA at which Brooks and MacWhinney found the most consistent effect of phonological priming across age groups and prime types. Overall, the experiments by Seiger-Gardner, Brooks, and Schwartz show that the PWI task has the potential to illuminate differences across clinical groups in language-formulation abilities, while the inconsistencies across these studies call for more experimental work on this task.

Speech Sound Disorder

The studies discussed thus far focused on TD children and those with SLI. As described earlier, children with Speech Sound Disorder (SSD) produce speech sounds less accurately than children with typical development (or those with SLI), in the absence of an obvious predisposing condition. There is little consensus on the underlying causes of SSD in children. Given that the most easily observable difficulties of children with SSD appear to be localized to their phonological output, it is logical to hypothesize that their difficulties might be the consequence of difficulties with the basic cognitive-linguistic processes that support the phonological stage of language production. A logical question, then, is to ask whether SSD is associated with a deficit in phonological encoding processes during language production.

To explore this possibility, the current study compares the performance of children with SSD to children with TD on a PWI task similar to that used by Brooks and MacWhinney (2000). Put differently, this article investigates whether a communication disorder that is often described as resulting from a specific impairment to the phonological stage of production can be better understood by examining performance on a task that is conventionally described as measuring the phonological stage of production. This study is part of a larger study that used the Levelt et al. (1999) model to examine the locus of impairment in children with SSD, other reports from which can be found in Munson, Baylis, Krause, and Yim (2010) and Yim and Munson (2012).

Based on previous findings, we hypothesize that both groups of children will name pictures more slowly when IWs are phonologically unrelated to the picture's name, compared to when they are phonologically related to the picture's name (i.e., when they are identical, onset-related, or rime-related). Our primary hypothesis is that for children with SSD, picture naming response times will resemble those of younger, typically developing children. That is, we predict that they will be facilitated more from hearing rime-related IWs than will those of typically developing age matches. We predict no group differences in the facilitative effect of hearing IWs that share onsets. In addition, we include the full range of SOAs and interfering-stimulus types found in Brooks and MacWhinney (2000), to ensure that their findings are replicable in our TD population.

Method

The current study is part of a larger study examining cognitive-linguistic processing by children with SSD and typically developing peers (Munson, Baylis, Krause, & Yim, 2010). The entire protocol that children participated in included standardized and nonstandardized tests of speech, language, and hearing, as well as three experimental tasks, including the picture-word interference task reported in this paper. Task order was randomized across children, with approximately equal numbers of children participating in different task orders.

Participants

The participants in the current study were 36 children, aged three to seven years. Participants for the study were recruited from the larger Minneapolis and St. Paul, Minnesota area. Recruitment took the form of word of mouth, flyers posted in the Speech-Language-Hearing-Sciences Department at the University of Minnesota – Twin Cities, local speech and language clinics (including local elementary school districts), local community organizations, and local child-care centers. All children were reported by their parents to be native, monolingual English speakers, and had no history of hearing impairment, psychosocial disorder such as autism, neuromotor disorder, or intellectual impairment. Eighteen of the children had a prior diagnosis of SSD, made either through their school district or a private clinic. The other 18 children had no history of speech, language, or hearing disorders. Each child with SSD was matched to a child with TD for age +/− 2 months. The children were a subset of the participants in the larger study. The 18 children with SSD comprised all of the children with a diagnosis of SSD who were able to complete the experimental task in this paper (out of the 25 children with SSD who enrolled in the larger project), and 18 age-matched children without SSD.

Children completed a series of standardized and nonstandardized assessments to measure their speech, language, hearing, and nonverbal IQ skills. The Sounds-in-Words subtest of the *Goldman-Fristoe Test of Articulation-2* (GFTA-2, Goldman & Fristoe, 2000) was used to measure speech-production accuracy. Two scores were extracted from this measure. The first was the conventional percentile rank score, in which the accuracy of the production of selected sounds in words was tallied and a percentile rank was determined based on the total number of errors. Second, each child's total percent phonemes correctly produced (PPC) was calculated. PPC scores were based on broad phonetic transcriptions of children's

productions of entire words, made by a trained phonetic transcriber who was blind to children's group membership. Each child's score was rationalized arcsine transformed for statistical analyses. This latter score is referred to henceforth as *GFTA-2 Total PPC*, and served two purposes. First, it was intended to be a finer-grained assessment of speech-production accuracy than is reflected by the GFTA-2 percentile rank, which takes into account only consonant-production accuracy for selected sounds. Second, its wider variability among participants made it a more appropriate measure for use in regression analyses to examine the association between speech-production accuracy and measures of lexical access. Given the inclusionary criteria for this study, the distribution of GFTA-2 percentile ranks was bimodal, and thus was not an appropriate variable for use in multiple regression. In contrast, the GFTA-2 Total PPC scores more closely approached a normal distribution, and were thus more amenable to use in regression.

The *Khan-Lewis Phonological Analysis-2* (KLPA-2, Khan & Lewis, 2002) was also used to examine whether children's errors followed common patterns. The *Peabody Picture Vocabulary Test-III* (PPVT-III, Dunn & Dunn, 1997) and the *Expressive Vocabulary Test* (EVT, Williams, 1997) measured children's receptive and expressive vocabulary, respectively. The matrices subtest of the *Kaufman Brief Intelligence Test* (K-BIT, Kaufman & Kaufman, 1990) was used as a screening measure of non-verbal intelligence. Normative scores on the K-BIT are provided for children over 4 years of age. Fourteen children in each group were old enough to complete this measure.

In addition to these standardized tests, all children completed three non-standard measures. The first was a speech discrimination task, described in detail in Baylis, Munson and Moller (2008), in which children identified words presented in minimal pairs. Forty-one sets of minimal pairs of pictures were selected. Stimulus pictures were taken from the corpus described in Bates et al. (2004). These word pairs featured initial and final position phoneme contrasts, such as *boat-goat*, which differ in the place of articulation of the initial consonant, and *pan-man*, which differ in the voicing and nasality of the initial consonant. Stimuli were produced by an adult male and were recorded for audio presentation. Stimuli used in the task were determined to be 100% intelligible to a group of naïve adult listeners. The child was seated at a table as single auditory stimulus items were presented at 65dB HL using a Dell Latitude D600 laptop, E-Prime software, and two Audix speakers (Model PH5). As each auditory item was presented, a pair of black-and-white picture cards was shown to the child, who was asked to point to the correct response. Responses were scored as correct or incorrect. Overall percentage correct was calculated for each child, and was rationalized arcsine transformed for statistical analyses.

The second nonstandardized measure was a baseline test of nonlinguistic speed of processing. This was an auditory-verbal response time (henceforth AVRT) task. During this task, children were instructed to say the word *yes* as soon as they heard a 100 ms, 1000 Hz pure tone presented through headphones. Each child's average nonlinguistic response time in ms was calculated, excluding trials that occurred greater than 5 s after the tone, and ones that were greater than 2.5 standard deviations above or below the child's mean nonlinguistic RT. This was intended to exclude response times that may have resulted from inattention or impulsivity.

The third non-standardized task was the PWI task described in greater detail below. In addition to the standardized and non-standardized measures of language and processing speed, several other screening-type assessments were conducted. Children participated in a diadochokinetic rate task to determine possible group differences in speech-motor control. Children repeated strings of the syllables [pʌ], [tʌ], [kʌ], and [pʌtʌkʌ] (modeled by the experimenter) as fast as they could. Children's speech production rate in syllables per second for contiguous, correctly produced sequences (including ones produced with habitual place and voicing errors by children with SSD) was hand-measured from digitized waveforms for each of these sequences. These were averaged together for analysis. Children were also given a nonstandardized examination of intraoral structures. There were no group differences in speech structures (e.g., dental-occlusal differences such as missing teeth, submucous cleft palate, or other structural anomalies).

All children completed pure-tone hearing screenings, in which they identified 0.5, 1, 2, and 4 kHz pure tones presented bilaterally at 20 dB HL. Sixteen children in each group passed the screening at this criterion; those who failed it detected the tones only when they were presented at 25 dB. These children's data were not qualitatively different from the rest of the children, and their data was thus retained in all analyses. Tympanometry screenings were also administered. Two children in both groups demonstrated abnormal results. Again, the findings were similar regardless of whether these individuals' data were included. All of the tasks described in this section were completed prior to participating in the PWI experiment.

PWI Stimuli

The target words were 11 black and white line drawings chosen from a collection of pictures described by Snodgrass and Vanderwort (1980). Two of these pictures (*cow* and *tree*) were used in the task-training procedure described below. The pictures used in this experiment were chosen because their names were Consonant-Vowel-Consonant (CVC) sequences; because they represented objects that were likely to be familiar to children; and because they were found by Bates et al. (2004) to have a high probability of being named uniformly. They also contained only the earliest-acquired phonemes of English /h/, /w/, /p/, /b/, /m/, /n/, /g/, /k/, /t/, /d/, /f/.

In this study there were five types of interfering stimuli, including four IWs and a nonlinguistic stimulus. *Identical* IWs were identical to the picture's name (picture: *cow*, interfering word: *cow*). *Onset-related* IWs shared the initial consonant with the picture's name (picture: *cow*, word: *cat).* *Rime-related* IWs shared the final VC sequence with the picture's name (picture: *cow*, word: *how*). *Phonologically unrelated* words shared no phonemes with the picture's name (picture: *cow*, word: *log*). In the *nonlinguistic auditory* condition, pictures were presented concurrent with a warble tone. This last condition was intended as a neutral baseline condition, analogous to the *go* condition in Brooks and MacWhinney (2000). A list of the IWs and target words used in this experiment can be found in Table 2. This table reports the average frequency of use and phonological neighborhood density of the IWs, using the values described in Pisoni, Luce, Nusbaum, and Slowiaczek (1985). Two Kruskal-Willis tests showed that the values did not differ significantly as a function of condition ($\chi^2_{[df=3,n=36]} = 1.44$, $p = 0.70$ for word frequency,

$\chi^2_{[df=3,n=36]} = 0.29$, $p = 0.96$ for neighborhood density). However, it should be pointed out that the neighborhood densities of the interfering words overall was relatively high, with the number of single-phoneme edit-distance neighbors ranging from 7 to 33. We return to this in the discussion.

The interfering words were all produced by a single adult male speaker of the local dialect. The durations of all of the interfering words were normalized so that each one was 480 ms long (roughly the average duration of the unnormalized stimuli) to prevent the possibility that naming latency differences across conditions would be spuriously influenced by differences in IW duration. This normalization was done using the PSOLA algorithm in the Praat signal-processing program (Boersma & Weenink, 2004), and required lengthening or shortening by no more than 10% of the original duration. The intelligibility of the resulting stimuli was assessed by having a group of three native speakers of English listen to the words and report what they were; they were all perceived 100% accurately, and were judged qualitatively to sound highly natural. All stimuli were equated for RMS amplitude.

Procedure

The picture-word interference task comprised three phases. During the *training* phase, the children were taught the names of eleven pictures. Nine of these pictures were the experimental target stimuli and two were practice stimuli. These pictures were presented to children in a laminated, spiral-bound book. After the pictures' names were trained, the experimenter taught the children the general routine of the experiment. The experimenter asked the children to name the pictures from the book while the experimenter said an interfering word using live voice. The purpose of this phase was to train the children on the picture's names, and to familiarize the children with the process of naming a picture while an interfering word was present.

In the next two phases, stimuli were presented from a computer. This was a Dell Latitude D600 computer attached to high-quality headphones. During the *practice* phase, three of the pictures were presented on a computer screen, and children heard the interfering words through headphones. The interfering stimuli were played at a level of approximately 65 dB, as calibrated prior to the experiment. Three different SOAs (−150 ms, 0 ms, and +150 ms) were used for the onset of the interfering stimulus relative to the onset of the picture on the computer screen. That is, the interfering stimulus was played either 150 ms before, concurrent with, or 150 ms after the onset of the presentation of the picture. The practice phase consisted of 45 stimulus-IW combinations.

During the *experimental* phase, the procedure was the same as for the practice phase but the target words changed to those listed in Table 2. Each of the nine pictures was named five times at each of the three SOA's, for a total of 135 stimuli. These IW-SOA combinations were divided randomly into three blocks of 45 stimuli. Blocks of stimuli were separated by brief breaks. Within blocks, stimuli were fully randomized. The experiment was created and executed using the E-prime experiment-management program.

The for both the practice and experimental phase, participants were instructed to name the pictures as soon as they knew what the pictures were and to try not to let the words from the

headphones "trick them." The interfering words and the participants' responses were recorded onto separate tracks of a stereo recording (the IWs on one track recorded directly from the computer; the participant's response on the other track recorded directly from the microphone), using a Marantz CDR300 CD recorder. The children spoke into an AKG C420 condenser microphone.

Analysis

Three sets of dependent measures were examined. The first of these was response accuracy. Accuracy was measured for two reasons. First, response latencies were analyzed only for the stimuli that were repeated correctly or with a habitual phonological error. Second, an ancillary purpose of this study is to report on factors that affect the accuracy of children's responses in the PWI task. We explore whether the rate and types of errors differ as a function of SOA and IW type. These analyses are presented in Appendix A.

Children's responses were coded as one of five categories, described in Table 3. These categories were based on our observation of the most frequently occurring errors in the corpus of children's responses. As this table shows, responses were counted as correct even if they included a child's habitual phonological error. For children with SSD, a separate count was made of the words on which a child produced a habitual phonological error. Percent of errors for each category was tallied separately for each SOA by IW Type, for a total of 15 percentages for each error type. As Appendix A shows, the accuracy of children's responses was affected by and large by the same factors that affected the latencies of correct responses.

The latency from the onset of the interfering word to the onset of the children's production of the target word was hand-measured from the acoustic waveforms using the Praat signal-processing program (Boersma & Weenink, 2004). This was then corrected so that the resulting response time (RT) reflected the response time from the onset of the picture's presentation to the onset of the child's response, using the E-Prime generated log of the SOA that was presented on a given trial. This is shown in Figure 1, which presents a schematic of a single trial from the PWI experiment.

Statistical outliers, defined as responses occurring outside 2.5 standard deviations from the mean for each participant, were removed from the RT analysis. RTs for words produced with a disfluency, a task-related error, or an off-task response (using the descriptions from Table 3) were excluded from analysis. For each participant, fifteen average response times were calculated from the remaining data: one for each of the 5 IW conditions (identical, phonologically related onset, phonologically related rime, phonologically unrelated, and non-speech auditory) at each of the 3 asynchronies (−150 ms, 0 ms, +150 ms).

Finally, a third dependent measure, response duration, was measured for correctly named responses only. Duration was measured so that we could explore whether it was influenced by the same factors that affected response latencies. These results are reported in Appendix A. For the purposes of the duration analysis, responses were counted as correct if they matched the adult targets for the pictures' names, and if any errors that they had were restricted to substitutions or distortions.

Results

This section reports the results of analyses for response latency measures. Analyses of response accuracy and response duration are presented in Appendix A. In all ANOVAs, the conservative Huyn-Feldt correction for sphericity was used for all within-subjects factors and interactions.

The primary dependent measure that we examined was response latency relative to the neutral condition, which in this experiment was the nonlinguistic auditory condition. For each participant, we calculated the average RT in the nonlinguistic auditory condition. The RT for each trial in the other conditions was then calculated relative to their by-subject average RT in the nonlinguistic auditory condition. Negative values indicated faster RTs than the nonlinguistic auditory condition, and positive values indicated slower RTs. This is the dependent measure that was presented in a number of previous PWI studies, including Brooks and MacWhinney (2000). The purpose of this analysis was to examine the influence of IW type while controlling for the presence of an interfering stimulus. For this analysis, 12 average normalized RTs were calculated per subject: one for each of the four IW conditions at each of the three SOAs. For reference, the raw RTs are shown in Table 4, separated by group, IW type, and SOA.

A three-factor mixed-model ANOVA was used to examine normalized RTs, with SOA (3 levels: −150 ms, 0 ms, +150 ms) and IW type (4 levels: onset-related, rime-related, identical, unrelated) as the within-subjects factor, and group (2 levels: SSD, TD) as the between-subjects factor. There was a significant main effect of SOA, $F[1.99,67.74] = 4.466$, $p = 0.015$, partial $\eta^2 = 0.11$. There was also a significant main effect of IW type, $F[2.58,87.88] = 47.152$, $p < 0.001$ partial $\eta^2 = 0.58$. These were mediated by a significant interaction between these two factors, $F[5.16,175.49] = 3.682$, $p = 0.003$, partial $\eta^2 = 0.10$. Finally, there was a significant main effect of group, $F[1,34] = 5.817$, $p = 0.021$, partial $\eta^2 = 0.15$. None of the other interactions was significant.

The effects can be seen by comparing the data in Figures 2 and 3. As these figures show, children with SSD had lower (faster) normalized RTs than TD children. That is, their RTs in the four real-word IW conditions showed either less inhibition (for the positive RTs) or greater facilitation (for the negative RTs). The average normalized RTs over the 12 conditions was −25 ms for children with SSD, and 80 ms for the TD children. The complex interaction between SOA and IW type can be seen by qualitatively examining the data in Figures 2 and 3. First, the RTs for the unrelated IWs remain relatively steady across the three SOAs. In contrast, the RTs for the three other IW types decrease—i.e., show greater facilitation—in the +150 ms SOA. This decrease is particularly marked for the onset-related IWs. Three post-hoc two-factor ANOVAs examined the influence of IW type on RTs at the three SOAs. All three ANOVAs showed a significant main effect of SOA, $F[3,32] = \{19.45–27.50\}$, all p's < 0.001. At all three SOAs, identical and rime-related IWs were associated with faster RTs than unrelated IWs. However, it was only at the +150 ms SOA that the onset-related IW was associated with faster SOAs than the unrelated condition.

Notably, group did not interact with SOA or with IW type. This was contrary to our hypothesis that the children with SSD would show greater rime-related facilitation than TD children. One complication of this analysis is that the children in this study had a relatively large age range, and included children who were younger than those who had been examined previously. Indeed, half of the children in this study were younger than 5;0, which is the youngest age that has been examined in prior studies.

To examine whether age mediated these findings, we conducted a second analysis with age as a two-level categorical factor. The younger group comprised 18 children (9 pairs of SSD-TD matches) with an average age of 4;0, and the older group comprised 18 children with an average age of 6;2. We then conducted a three-factor mixed-model ANOVA similar to the one described above, but with age group as an additional between-subjects factor. In this ANOVA, age group did not interact with clinical group, either in a two-way interaction or in any higher-order interactions. The only additional interaction of note was that between age group and IW type, $F[2.94, 94.02] = 5.694$, $p = 0.001$ partial $\eta^2 = 0.15$. Averaged across SOAs, younger children showed larger facilitation and inhibition effects than did older children (for onset-related IWs: $M_{younger} = 262$ ms (SEM = 249 ms), $M_{older} = 140$ ms (SEM = 152 ms); for rime-related IWs: $M_{younger} = 28$ ms (SEM = 180 ms), $M_{older} = -31$ (SEM = 221 ms); for identical IWs: $M_{younger} = -376$ ms (SEM = 225 ms), $M_{older} = -214$ (SEM = 233 ms); for unrelated IWs: $M_{younger} = 304$ ms (SEM = 253 ms), $M_{older} = 118$ ms (SEM = 191 ms). We return to this topic in the discussion.

Regression Analyses

The analyses presented thus far suggest that children with SSD do not differ from children with TD in the predicted direction, as they did not show greater facilitation from hearing rime-related IWs than will did typically developing age-matched children. This section presents regression analyses examining predictors of individual children's production. As described earlier, the sample of children with SSD in this study differed from children with TD not only in speech-production accuracy, but in expressive vocabulary size and in phoneme discrimination. The purposes of these analyses were to (a) determine whether expressive vocabulary and discrimination ability mediated group differences in the PWI task, and (b) to examine predictors of performance on the PWI task more generally.

The first regression analysis examined whether the main effect of group in the ANOVA examining response times relative to the nonlinguistic condition was due to the group differences in speech-production accuracy, or to the group differences in vocabulary size. To examine this, average nonlinguistic-normalized RTs across the 3 SOAs and 4 IW types were calculated for each subject. This served as the dependent measure in a hierarchical multiple regression. In the first run, a conservative variable-entry order was used, in which age was forced as the first variable, followed by a block in which PPVT and EVT raw score and standard score were entered, as well as discrimination accuracy. This was followed by a third block, in which GFTA Total PPC was entered if it accounted for a significant proportion of variance beyond the variables entered in the previous blocks. This regression provided the strictest possible test of the hypothesis that the group differences were an artifact of group differences in perception or in vocabulary size. In this regression, only

GFTA Total PPC accounted for a significant proportion of variance in average nonlinguistic-normalized RT, $F[1,28] = 8.636$, $p = 0.007$, $R^2 = 22.1\%$, standardized $\beta = 0.75$. However, the full regression model was not significant, F[7,28] = 1.57, p > 0.05. A second run used a less conservative variable order, in which age was forced as the first variable, but all other variables were entered in a stepwise fashion if they accounted for a significant increase in the proportion of variance accounted for ($\alpha < 0.05$) beyond those variable(s) entered on the previous step(s). In this regression, age did not account for a significant proportion of variance. On the second step, GFTA Total PPC accounted for 11.7% of the variance in the dependent measure, $F[1,33] = 4.379$, $p = 0.044$, standardized $\beta = 0.619$. Children with more accurate speech production showed greater overall interference in the PWI task.

The second set of regressions examined predictors of average raw RTs (i.e., those RTs shown in Table 4). These analyses were motivated by the fact that relatively little work has examined response times in children with SSD, and hence it is unclear whether RTs in this population are subject to the same predictors as are RTs in other, better-studied populations, like children with SLI and typically developing children. In these regressions, RTs were averaged across the five IW types at each SOA. These were the dependent measures in three hierarchical multiple regression analyses. In each regression, age and nonlinguistic RT were forced in the first block. In the second block, EVT and PPVT raw scores, GFTA Total PPC, and Discrimination Accuracy were entered if they accounted for a significant proportion of variance in the dependent measure beyond what was accounted for by the variable(s) entered on the previous step(s). The results of these regressions are shown in Table 5. As this table shows, age consistently predicted a significant proportion of variance in RTs. As expected, older children had faster naming times than younger ones. At the −150 ms delay interval, no other variables accounted for additional variance in the dependent measure. At the 0 ms and +150 ms delay intervals, however, both vocabulary-size measures predicted significant variance in naming latency. In both cases, children with larger expressive vocabularies had faster naming speeds than children with smaller ones. Interestingly, however, children with larger-sized receptive vocabularies had *slower* naming speeds than children with smaller-sized ones. We return to this asymmetry in the discussion. In none of these regressions did GFTA, Discrimination Accuracy, or DDK Rate account for a significant proportion of variance in RTs.

Discussion

This article reported on the results of a picture-word interference experiment, designed to examine whether children with Speech Sound Disorder have less efficient phonological encoding mechanisms than children with typical phonological development. The findings presented in this paper do not support the hypothesis that such differences exist. As in previous research, both groups of children named pictures more quickly when they were presented with a phonologically similar interfering word (IW) than when they were presented with a phonologically unrelated IW. This effect was strongest when the IW was presented 150 ms after the picture was presented, presumably because this coincided with the point in time when children were beginning to formulate plans for the production of the pictures' names. Moreover, there was no overall group difference in the influence of identical IWs on task performance. This finding confirms that both groups had a similar

propensity to be affected by the presence of interfering auditory stimuli. That is, the lack of group differences in the influence of interfering words suggests that group differences in speech sound production are unrelated to cross-modal integration of auditory and visual stimuli. Critically, we did not find evidence that children with SSD had greater facilitation of rime-related IWs than children with TD. Such a finding would have suggested that the phonological encoding of children with SSD is like that of younger, typically developing children. The only group difference was that children with SSD had lower normalized RTs overall, suggesting that they experienced both less inhibition and greater facilitation from IWs than did the children with TD.

The findings in this paper have clear implications for research on the PWI task. Perhaps the most interesting implication for future research comes from the finding that expressive and receptive vocabulary size affects performance on the PWI task differentially. This suggests that having a large-sized expressive vocabulary is associated with faster naming speeds—a finding that is strongly consistent with much previous research (e.g., Lahey & Edwards, 1996). The more surprising finding is that larger-sized receptive vocabularies are associated with slower naming speeds on the PWI task. Given that receptive and expressive vocabulary measures both reflect the size of individuals' mental lexicons, this finding is, at first glance, unexpected. We might expect that naming latencies would be inversely correlated with both measures.

One possible explanation for this finding relates to the fact that the IWs in this study were almost uniformly high in phonological neighborhood density. In word-recognition studies, stimuli from dense phonological neighborhoods are recognized less accurately than words from sparse neighborhoods. This finding can be interpreted within models of spoken-word recognition (e.g., Luce & Pisoni, 1998) as evidence that stimuli activate words in the same phonological neighborhood, even when those words are not relevant for the recognition task. This activation is so automatic that it occurs across the lexicons of bilingual individuals (Spivey & Marian, 1999). The target words can only be recognized accurately once the phonologically related competing candidates are suppressed.

The association between receptive vocabulary size and naming speed in this study may have been, in part, a consequence of the density of the IWs that were used. That is, the lack of control over the density of the IWs led people with larger-sized receptive vocabularies to be unexpectedly disproportionately disadvantaged by the fact that the IWs activated more competing words in the listeners' receptive lexica than low-density words would have. This conjecture warrants a rigorous prospective examination, using IWs that vary systematically in their phonological neighborhood density. Moreover, it suggests that lexical characteristics of IWs might explain apparently discrepant findings among previous studies using this paradigm.

Finally, one word of caution needs to be made about the design of this study. Unlike in some previous studies (including Brooke & MacWhinney, 2000 and Seiger-Gardner & Brooks, 2008), IWs that represented the picture's names also served as other types of IWs for other target words. For example, the spoken word *bed* served as the identical IW for the target word *bed* but as the unrelated IW for the pictures *cat* and *knife* and as the onset-related IW

for the picture of *boat*. As shown by Hanaeur and Brooks (2005), the interference provided by a spoken IW increases when it is part of the response set, i.e., when it is one of the picture's names. The fact that this variable was not balanced across the experiment might have led to less even patterns of interference than might have been found if all IWs were not members of the response set. This slightly limits the extent to which the results can be compared to these previous studies.

The larger issue addressed by this study was to illuminate the underlying causes of SSD in children by examining how these children perform on a task that purports to measure on-line language formulation processes. The task was chosen because it was designed to test the explicit model of production outlined in Levelt et al. (1999). However, that model was designed to account for language production in adults, in which errors occur at very low rates and generally do not resemble those of children with SSD (i.e., they are more likely to involve repetitions or transpositions of phonemes in adjacent words, rather than phonological simplifications). Perhaps a different model is needed as a jumping-off point for understanding SSD. It is here that research on the nature of SSD suffers, as there are relatively few models of production to use as the basis for a program of research on SSD. One promising recent proposal is from Redford (2015), who outlines components of what a developmentally sensitive model of production would need to include.

In the absence of useful theoretical models of production, studies of the nature of SSD have relied on a mix of tasks derived from different theoretical frameworks to examine the nature of this disorder (Waring & Knight, 2013). One example of this is Munson, Edwards, and Beckman's (2005) study of SSD, which was informed by the model of phonological knowledge presented in Munson, Edwards, and Beckman, (2011), which itself is very similar to the model described by Dodd, Leahy, and Gambly (1989). The Munson et al. and Dodd et al. models emphasize the sensory domains in which phonological generalizations are made. The model that inspired this study emphasizes the stages in production, emphasizing the processing of relatively abstract phonological elements.

Another tactic for studying the nature of SSD is to use purely statistical exploratory techniques to examine the factors that underlie SSD, a possibility suggested by Munson, Bjorum, and Windsor (2001) and implemented by Lewis et al. (2006) and Vick et al. (2015). However, even the most methodologically rigorous work in this vein is limited by the tasks chosen as input to the model. That is, a model cannot uncover a skill area that is not sampled by the tasks used as input to the model. Moreover, any complete model of adults' production processes must also be able to model pronunciation variation seen in children. Ultimately, we agree with Redford (2015) that experimental studies of children's speech production, including those of children with SSD, must be central in the formulation of any model of production. Such a model could potentially provide a substantial improvement in the taxonomies of SSD, as well as in the assessment and treatment of this disorder.

Acknowledgments

the audience at the Tenth Conference on Laboratory Phonology and the 2006 Symposium for Research on Child Language Disorders for useful comments on this work as it was in progress.

References

Bates E, D'Amico S, Jacobsen T, Székely A, Andonova E, Devescovi A, Herron D, Lu CC, Pechmann T, Pléh T, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Hsu J, Iyer G, Kohnert K, Behotcheva T, Orozco-Figueroa A, Tzeng A, Tzeng O. Timed picture naming in four languages. Psychonomic Bulletin and Review. 2004; 10:344–380.

Baylis A, Munson B, Moller K. Factors affecting articulation skills in children with velocardiofacial syndrome and children with cleft palate or velopharyngeal dysfunction: A preliminary report. Cleft Palate Craniofacial Journal. 2008; 45:193–207. [PubMed: 18333642]

Boersma, P., Weenick, D. Pratt 3.9.15 [Computer software]. Amsterdam: Institute of Phonetic Sciences; 2001.

Brooks P, MacWhinney B. Phonological priming in children's picture naming. Journal of Child Language. 2000; 27:335–366. [PubMed: 10967891]

Crosbie S, Holm A, Dodd B. Cognitive flexibility in children with and without speech disorder. Child Language Teaching and Therapy. 2009; 25:250–270.

Dodd B, Leahy J, Hambly G. Phonological disorders in children: underlying cognitive deficits. British Journal of Developmental Psychology. 1989; 5:55–71.

Dodd B, McIntosh B. Two-year-old phonology: impact of input, motor and cognitive abilities on development. Journal of Child Language. 2010; 37:1027–1046. [PubMed: 19961658]

Dunn, L., Dunn, L. Peabody Picture Vocabulary Test-III. Circle Pines, MN: American Guidance Service; 1997.

Edwards J. Compensatory speech motor abilities in normal and phonologically disordered children. Journal of Phonetics. 1992; 20:189–207.

Edwards J, Fox RA, Rogers CL. Final consonant discrimination in children: Effects of Speech Sound Disorder, vocabulary size, and phonetic inventory size. Journal of Speech, Language, and Hearing Research. 2002; 45(2):231–242.

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

Hanauer J, Brooks P. Contributions of response set and semantic relatedness to cross-modal Stroop-like picture–word interference in children and adults. Journal of Experimental Child Psychology. 2005; 90:21–47. [PubMed: 15596075]

Jerger S, Martin R, Damian M. Semantic and phonological influences on picture naming by children and teenagers. Journal of Memory and Language. 2002; 47:229–249.

Kahn, L., Lewis, N. Kahn-Lewis Phonological Analysis-2. Circle Pines, MN: American Guidance Services; 2002.

Kail R. Processing time declines exponentially during childhood and adolescence. Developmental Psychology. 1991; 27:259–66.

Kaufman, AS., Kaufman, NL. Kaufman Brief Intelligence Test. Circle Pines, MN: American Guidance Service; 1990.

Lahey M, Edwards J. Why do children with specific language impairment name pictures more slowly than their peers? Journal of Speech and Hearing Research. 1996; 39:1081–1098. [PubMed: 8898260]

Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. International Journal of Language and Communication Disorders. 2000; 35:165–188. [PubMed: 10912250]

Levelt W, Roelofs A, Meyer A. A theory of lexical access in speech production. Behavioral and Brain Sciences. 1999; 22:1–75. [PubMed: 11301520]

Lewis B, Freebairn L, Hansen A, Stein C, Shriberg L, Iyengar S, Taylor H. Dimensions of early speech sound disorders: A factor analytic study. Journal of Communication Disorders. 2006; 39:139–157. [PubMed: 16386753]

Lewis B, Freebairn L, Tag J, Ciesla A, Iyengar S, Stein C, Taylor H. Adolescent outcomes of children with early speech sound disorders with and without language impairment. American Journal of Speech-Language Pathology. 2015; 24:150–163. [PubMed: 25569242]

Luce P, Pisoni D. Recognizing spoken words: the neighborhood activation model. Ear and Hearing. 1998; 19:1–36. [PubMed: 9504270]

Munson, B., Baylis, AL., Krause, MO., Yim, D. Representation and Access in Phonological Impairment. In: Fougeron, C.Kühnert, B.D'Imperio, M., Vallée, N., editors. Laboratory Phonology. Vol. 10. Berlin: Mouton de Gruyter; 2010. p. 381-404.

Munson B, Bjorum E, Windsor J. Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. Journal of Speech, Language, and Hearing Research. 2003; 46:189–202.

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

Munson, B., Beckman, ME., Edwards, J. Phonological representations in language acquisition: climbing the ladder of abstraction. In: Cohn, A.Fougeron, C., Huffman, M., editors. Oxford Handbook in Laboratory Phonology. Oxford: Oxford University Press; 2011. p. 288-309.

Pisoni D, Nusbaum H, Luce P, Slowiaczek L. Speech perception, word recognition, and the structure of the lexicon. Speech Communication. 1985; 4:75–95. [PubMed: 23226910]

Preston JL, Hull M, Edwards ML. Preschool speech error patterns predict articulation and phonological awareness outcomes in children with histories of speech sound disorders. American Journal of Speech-Language Pathology. 2013; 22:173–184. [PubMed: 23184137]

Redford M. Unifying speech and language in a developmentally sensitive model of production. Journal of Phonetics. 2015; doi: 10.1016/j.wocn.2015.06.006

Rvachew S, Grawburg M. Correlates of phonological awareness in preschoolers with speech sound disorders. Journal of Speech, Language, and Hearing Research. 2006; 49:74–87.

Rvachew S, Jamieson D. Perception of voiceless fricatives by children with functional articulation disorder. Journal of Speech and Hearing Disorders. 1989; 54:193–208. [PubMed: 2709838]

Schriefers H, Meyer A, Levelt W. Exploring the time course of lexical access in language production: picture-word interference studies. Journal of Memory and Language. 1990; 29:86–102.

Seiger-Gardner L, Brooks P. Effects of onset- and rhyme-related distractors on phonological processing in children with specific language impairment. Journal of Speech, Language, and Hearing Research. 2008; 51:1263–1281.

Seiger-Gardner L, Schwartz R. Lexical access in children with and without specific language impairment: a cross-modal picture–word interference study. International Journal of Language and Communication Disorders. 2008; 43:528–551. [PubMed: 22612630]

Shriberg LD, Tomblin JB, McSweeny JL. Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. Journal of Speech, Language, and Hearing Research. 1999; 42:1461–1481.

Schwartz M, Dell G, Martin N, Gahlo S, Sobel P. A case-series test of the interactive two-step model of lexical access: evidence from picture naming. Journal of Memory and Language. 2006; 54:228–264.

Snodgrass J, Vanderwort M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. Journal of Experimental Psychology: Human Learning and Memory. 1980; 6:174–215. [PubMed: 7373248]

Spivey M, Marian V. Cross-talk between native and second languages: partial activation of an irrelevant lexicon. Psychological Science. 1999; 10:281–284.

Starreveld P, La Heij W. Time-course analysis of semantic and orthographic context effects in picture naming. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1996; 22:896–918.

Towne RL. Effects of mandibular stabilization on the diadochokinetic performance of children with Speech Sound Disorder. Journal of Phonetics. 1994; 22:317–332.

Vick J, Campbell T, Shriberg L, Green J, Truemper K, Rusiewicz H, Moore C. Data-driven subclassification of speech sound disorders in preschool children. Journal of Speech, Language, and Hearing Research. 2015; 57:2033–2050.

Waring R, Knight R. How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. International Journal of Language and Communication Disorders. 2013; 28:25–40.

Williams, K. Expressive Vocabulary Test. Circle Pines, MN: American Guidance Service; 1997.

Wren YE, Roulstone SE, Miller LL. Distinguishing groups of children with persistent speech disorder: Findings from a prospective population study. Logopedics Phoniatrics Vocology. 2012; 37:1–10.

Yim D, Munson B. Phonetic accuracy on a delayed picture-naming task by children with a phonological disorder. Korean Journal of Communication Disorders. 2012; 17:187–200.

Appendix A: Response Accuracy and Response Duration

This section presents statistics separately for each dependent measure: response accuracy, response time, and duration. In all ANOVAs, the conservative Huyn-Feldt correction for sphericity was used for all within-subjects factors and interactions.

Response Accuracy

Five separate three-factor mixed-model ANOVAs examined the influence of SOA (3 levels: −150 ms, 0 ms, +150 ms), IW Type (5 levels: nonlinguistic, onset-related, rime-related, identical, unrelated), and Group (2 levels: SSD, TD) on the five error types (overall response errors, habitual phonological errors, off-task errors, disfluent errors, and task-related errors). The first examined the proportion of accurate responses. Recall that this category encompassed responses both with and without a habitual phonological error. These are shown in Table A.1. In this ANOVA, only one significant effect was found, the main effect of IW type, $F[3.88,131.92] = 17.658$, $p < 0.001$, partial $\eta^2 = 0.34$. As Table A.1 shows, the highest accuracy rates (90%, pooled across groups and SOAs) was found for identical IWs, the lowest rate (74.3%) for onset-related IWs, and intermediate levels of accuracy (approximately 80%) for the remaining IW types. All pairwise differences were significant except three: nonlinguistic/rime-related, onset-related/unrelated, and rime-related/unrelated.

The second ANOVA, identical in design to the first, examined the proportion of target words with a habitual phonological error. In this ANOVA, the only significant main effect was of group, $F[1,34] = 14.432$, $p = 0.001$, partial $\eta^2 = 0.30$. Not surprisingly, children with SSD produced a higher proportion of words with a phonological error (22%, averaged across SOAs and IW types) than children with TD (6%).

The remaining three ANOVAs examined predictors of the other three error types. In the ANOVA examining off-task responses, a small but significant effect of IW was found, $F[3.31,117.71] = 2.908$, $p = 0.033$, partial $\eta^2 = 0.08$. This arose because there was a significantly higher rate of non-responses for onset-related IWs than for identical IWs. In the ANOVA examining the rate of disfluent responses, an effect of IW type was again found, $F[3.8,129.2] = 12.824$, $p < 0.001$, partial $\eta^2 = 0.27$. Here, the highest proportion of disfluent responses was found for onset-related and unrelated IWs (approximately 8% each, pooled across groups and SOAs); the lowest rate for identical IWs (1%), and intermediate rates for rime-related IW and nonlinguistic stimulus (approximately 4% each). All pairwise

differences were significant except those between onset-related and unrelated IWs, and between rime-related and nonlinguistic IWs.

In the ANOVA examining task-related errors, the main effect of SOA approached significance, F[1.82, 62.73] = 3.096, p = 0.056, partial η^2 = 0.08. This arose because there was a smaller rate of these (approximately 1%, pooled across groups and IW types) at the −150 ms delay interval than at the other two delay intervals (approximately 2% for each). There was also a significant main effect of IW type, F[3.64,123.77] = 4.431, p = 0.003, partial η^2 = 0.12. This arose because the proportion of errors in the identical IW condition was significantly smaller than those in the unrelated and rime-related IW conditions.

From this complex set of analyses we can extract three key findings. First, group differences in most types of error rates are minimal. The only significant main effect of group was found for the one category in which we would expect group differences: the rate of phonological errors. Moreover, the influence of SOA was also relatively minimal, affecting only the rate of off-task responses, with more off-task responses occurring when the IW was presented relatively later. In contrast, the principle factor affecting error rates was IW type. In general, the highest rate of accurate productions was found, not surprisingly, for the identical IW, where the IW likely served as a direct cue to the picture's name. The highest error rates were found in general for the unrelated IWs, though overall accuracy levels were also lower for words named in the onset-related IW conditions.

Response Duration

The final ANOVA examined response duration. The design of this ANOVA was again identical to those described earlier. The main effect of group did not achieve statistical significance at the a < 0.05 level, but did approach it, F[1,34] = 3.946, p = 0.055, partial η^2 = 0.10. Children with SSD produced words with shorter durations than children with TD. Moreover, there was a significant main effect of IW type, F[3.72, 126.56] = 5.385, p = 0.001, partial η^2 = 0.14. None of the interactions were significant. Post-hoc Bonferroni-corrected paired comparisons showed that the response durations were significantly shorter in the identical condition than in the rime-related condition (mean difference = 40 ms) and the onset-related condition (mean difference = 35 ms). No other pairwise differences were significant. This finding was somewhat surprising. As a follow-up analysis, we examined whether the longer production times in these conditions were due to selective lengthening of the onset (in the onset-related condition) or the rime (in the rime-related condition). Onset and rime durations were measured. For each word, the proportion of the total duration comprising the onset was calculated. For each subject, averages for this measure were taken separately by IW type and SOA. A three-factor mixed-model ANOVA found no effect of group, SOA, or IW type on this measure.

Another possible interpretation of this unexpected effect of IW type on duration is that the longer target-word durations in the phonologically related IW conditions reflected a trade-off between planning and execution times, as the phonologically related IWs had an overall facilitative effect on production latencies. However, this conjecture would fail to explain why the influence of IW type on duration did not interact with SOA, as it did on production

latencies. Moreover, it would fail to explain why the durations in the identical IW condition were not longer, as this was condition in which latencies were shortest. While the finding in the current manuscript demonstrates clearly that production latencies do not tell the whole picture about performance on the PWI task, further research is needed to clarify the nature of the association between IW type and production duration.

Table A.I

Picture-naming accuracy, either with or without a phonological simplification (e.g., [tæt] for kæt/)

SOA	IW Type	TD Mean	TD SD	SSD Mean	SSD SD
−150 ms	Nonlinguistic	87.0%	13.9%	79.6%	23.9%
	Unrelated	84.6%	12.7%	75.3%	20.4%
	Onset-Related	77.2%	18.5%	72.8%	16.3%
	Rime-Related	84.6%	13.3%	77.2%	24.8%
	Identical	92.6%	12.6%	88.3%	18.9%
0 ms	Nonlinguistic	88.9%	12.6%	78.4%	19.6%
	Unrelated	81.5%	21.6%	74.1%	19.1%
	Onset-Related	78.4%	17.7%	71.6%	19.5%
	Rime-Related	79.6%	17.6%	81.5%	16.2%
	Identical	88.3%	14.0%	90.1%	15.7%
+150 ms	Nonlinguistic	86.4%	13.0%	82.1%	19.9%
	Unrelated	79.0%	14.2%	75.3%	18.1%
	Onset-Related	75.9%	17.6%	69.8%	22.5%
	Rime-Related	87.0%	14.9%	80.9%	16.1%
	Identical	90.7%	14.4%	90.7%	10.9%
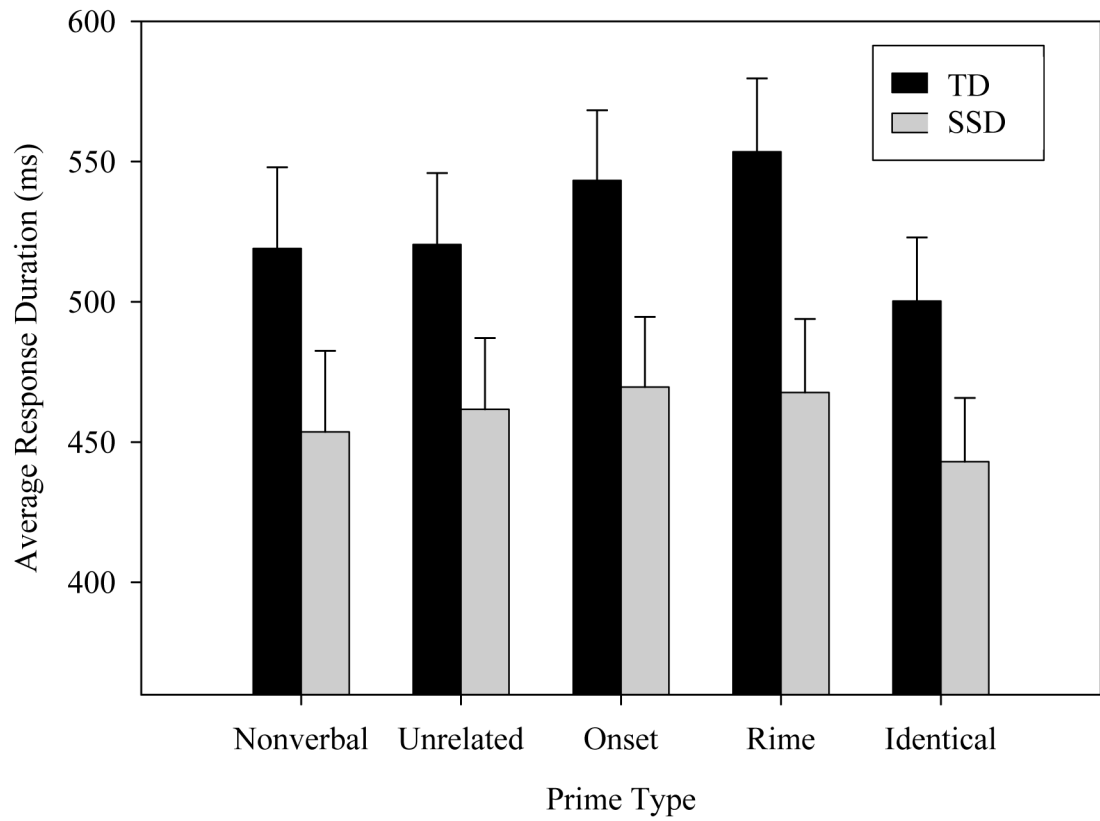
Figure A.1.
Average response durations for children with TD and children with SSD, separated by IW Type and averaged across SOA.

What this Paper Adds

There is no consensus on what causes some children to have Speech Sound Disorder (SSD, i.e., highly inaccurate speech production in the absence of any clear predisposing condition). The psycholinguistic model of language production formulated by Levelt and colleagues (Levelt, Roelofs, & Meyer, 1999) proposes that production occurs in multiple stages from intention to articulation. This paper uses this as a framework for examining the stage in on-line language production during which breakdowns might occur in children with SSD. We report the results of a cross-modal picture-naming task that has been used previously to study phonological encoding during language production. We find that children with SSD perform similarly to their typically developing (TD) peers on this task. This suggests that the speech errors of children with SSD are not the consequence of reduced efficiency of phonologically encoding words that have been accessed from memory. This paper shows how a well-specified, widely used model of language production can be used to study the nature of SSD in children.
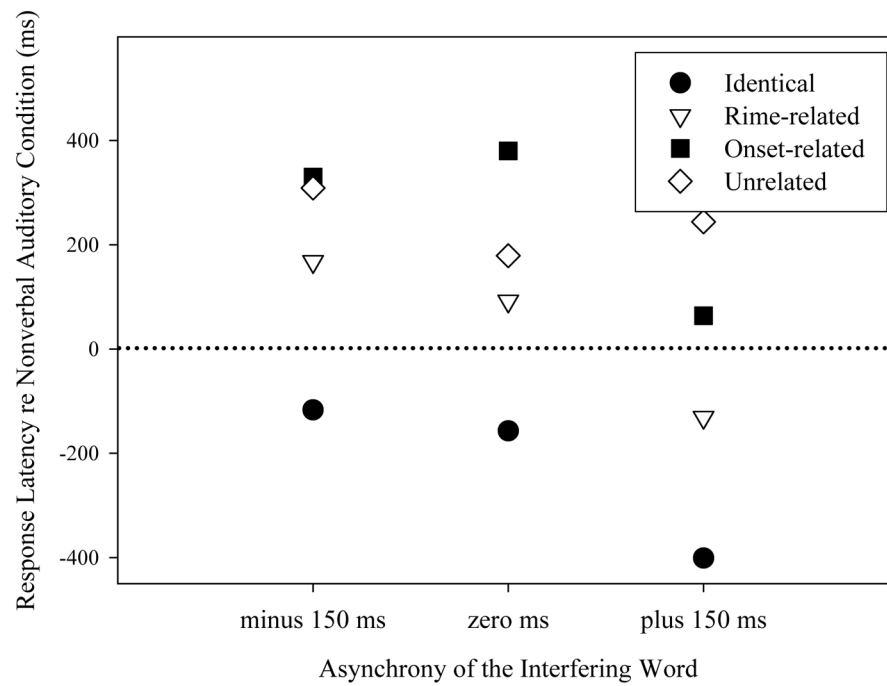
Picture (on continuously)

Interfering Word

Child's Response

SOA   +150 ms

RT   1200 ms
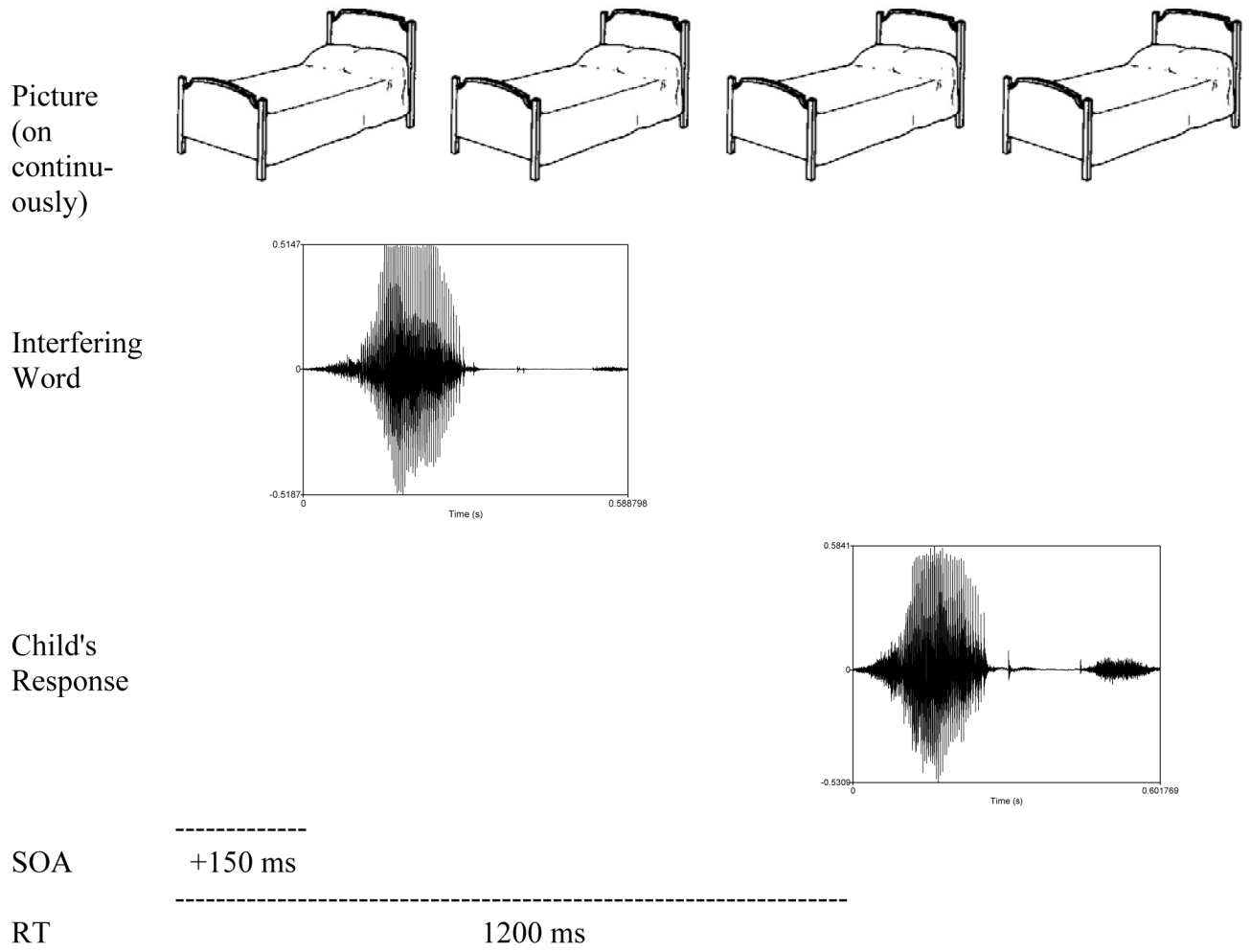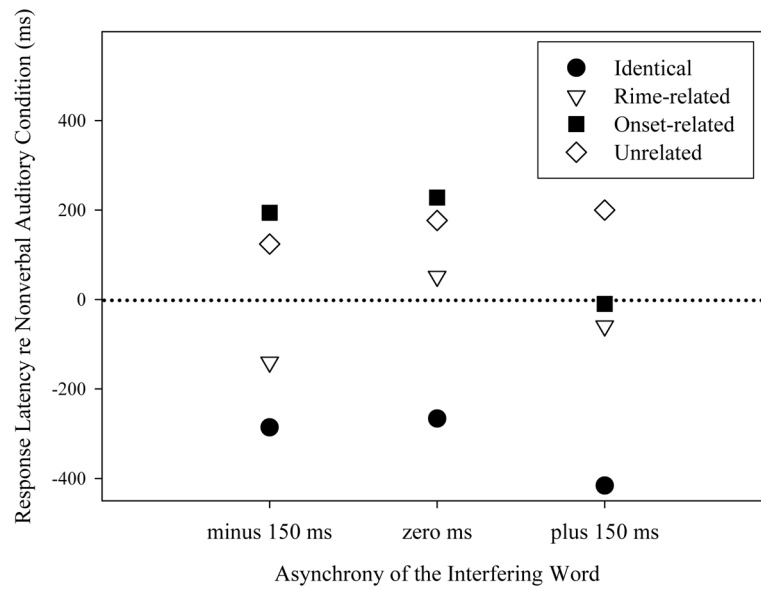
Figure 1.
Schematic illustration of the time-course of a single experimental trial.

Figure 2.
Average response latencies for the children with TD relative to the nonlinguistic auditory condition. RTs below the dotted line indicate performance faster than the nonlinguistic condition; those above the line indicate slower performance.

Figure 3.
Average response latencies for the children with SSD relative to the nonlinguistic auditory condition. RTs below the dotted line indicate performance faster than the nonlinguistic condition; those above the line indicate slower performance.

Table 1

Participant Characteristics

Measure	TD[a] Mean	SD	SSD[b] Mean	SD	Difference Test
Age (months)	62.1	15.6	61.2	16.6	$t[34] = -0.17, p > 0.10$
Percent Girls	77.8		38.9		$\chi^2_{[df=1,n=36]} = 5.6, p = 0.02$
GFTA-2[e,c]	65.6	18.5	8.4	5.1	$t[34] = 12.60, p < 0.001$
GFTA Total PPC[j]	95.4	4.4	70.5	17	$t[34] = 32.05, p < 0.001$[k]
KLPA[f,d]	107.8	5.8	77.7	18.6	$t[34] = 30.17, p < 0.001$
PPVT-III[g,d]	116.3	10.3	108.7	13.2	$t[34] = 1.91, p = 0.06$
PPVT-III[g,l]	4.3	0.4	4.4	0.2	$t[34] = 1.35, p > 0.10$
EVT[h,d]	118.8	11.6	106.6	10.4	$t[34] = 3.05, p = 0.004$
EVT[h,l]	4.2	0.2	4	0.3	$t[34] = 1.77, p = 0.09$
K-BIT[i,d]	108.4	12.7	110	10.4	$t[26] = 1.64, p > 0.05$
Average DDK[j]	4.7	0.7	4.4	0.9	$t[34] = 1.00, p > 0.05$
Nonlinguistic RT[k]	440.4	187.6	438.8	184.3	$t[34] = -0.03, p > 0.05$
Phoneme Discrimination[l]	105.1	8.6	96.4	12.1	$t[34] = 2.51, p = 0.02$

[a] Typically Developing,

[b] Phonologically Disordered,

[c] Percentile Rank,

[d] Standard Score (Mean=100, SD=15),

[e] *Goldman-Fristoe Test of Articulation-2* (Goldman & Fristoe, 2000),

[f] *Khan-Lewis Phonological Analysis* (Khan & Lewis, 2002),

[g] *Peabody Picture Vocabulary Test-III* (Dunn & Dunn, 1997),

[h] *Expressive Vocabulary Test* (Williams, 1997),

[i] *Kaufman Brief Intelligence Test* (Kaufmann & Kaufmann, 1990),

[I] Natural-Log Transformed Raw Score,

[j] Syllables per Second,

[k] in milliseconds,

[I] Rationalized Arcsine Transformed Units (Studebaker, 1985)

Table 2

Stimuli used in this experiment, not including the nonlinguistic auditory condition.

	Picture Name/Identical IW[1]	Onset-Related IW	Rime-Related IW	Unrelated IW
	bed	*bit*	*head*	*pin*
	boat	*bed*	*goat*	*make*
	cat	*keep*	*bat*	*bed*
	dog	*dad*	*hog*	*make*
	foot	*fish*	*put*	*keep*
	hat	*head*	*cat*	*knife*
	knife	*neat*	*wife*	*bed*
	pen	*pot*	*ten*	*make*
	wood	*wife*	*could*	*pen*
Average IW Log Frequency (SD)	2.9 (0.6)	2.9 (0.5)	2.9 (0.9)	3.2 (0.6)
Average IW Neighborhood Density (SD)	19.0 (9.8)	20.6 (6.6)	19.3 (8.4)	20.8 (5.5)

[1] IW: Interfering Word. Underlined IWs are those also used as target picture names.

Table 3

Error Types.

Category	Description	Example for target word *cat*, interfering word *bed*
Off-Task	No response, talking during the experimental trial	
Disfluency	Hesitation, false start	[k] [kæt]
Task-Related Error	Child says the interfering word rather than the picture's name, or says a different target word	"bed," "hat"
Correct Responses	Child says the target word. Children's habitual phonological errors were counted as correct	"cat", [tæt], [gæt]
Phonological Errors Only	The child says the target word with a non-habitual phonological error	[tæt], [gæt]

Table 4

Picture-naming latencies in ms (and standard errors of measurement [SEM]) for children with typical language development (TD) and children with Speech Sound Disorder (SSD), separated by interfering word (IW) type and stimulus onset asynchrony (SOA).

SOA	IW Type	Group			
		TD		SSD	
		Mean	SEM	Mean	SEM
−150 ms	Nonlinguistic	1341	67	1580	124
	Unrelated	1650	109	1705	101
	Onset-Related	1672	101	1775	111
	Rime-Related	1509	112	1440	81
	Identical	1223	62	1194	60
0 ms	Nonlinguistic	1494	69	1561	103
	Unrelated	1674	88	1738	116
	Onset-Related	1874	125	1789	121
	Rime-Related	1587	92	1613	96
	Identical	1337	84	1295	84
+150 ms	Nonlinguistic	1600	78	1643	112
	Unrelated	1844	103	1842	116
	Onset-Related	1908	98	1832	103
	Rime-Related	1713	93	1782	112
	Identical	1443	104	1426	101

Table 5

Regression analyses predicting average RTs from selected dependent measures. [ANOVA] for full model based on the average RTs at the −150 ms SOA; [ANOVA] for the full model based on average RTs at the 0 ms SOA; [ANOVA] for full model based on the +150 ms SOA.

Dependent Measure	Step	Independent Variable	$R^{2\,a}$	B^b	SE B^b	β^b
Average RT, −150 ms SOA	1	Age	29.0% **	−10.804	3.173	−0.498 **
	2	Nonlinguistic RT	3.4%	0.356	0.275	0.189
Average RT, 0 ms SOA	1	Age	21.7% **	−11.616	5.041	−0.511 *
	2	Nonlinguistic RT	3.8%	0.392	0.292	0.199
	3	PPVT raw score	5.9% *	733.666	282.942	0.656 **
	4	EVT raw score	8.3% *	−774.284	375.438	−0.534 *
Average RT, +150 ms SOA	1	Age	17.1% **	−11.809	4.724	−0.517 **
	2	Nonlinguistic RT	6.1%	0.518	0.274	0.262
	3	PPVT raw score	12.0% *	985.263	265.122	0.877 **
	4	EVT raw score	12.5% **	−954.680	351.793	−0.655 **

[a] Increase in R^2 over the model containing all previous steps,

[b] coefficients for the full model,

*
$p < 0.05$,

**
$p < 0.01$,