5-11-2013

Contextual Influences on Phonetic Categorization in Developmental Populations

Jean Alexandra Campbell
University of Connecticut - Storrs, jean.campbell@uconn.edu

Recommended Citation

Campbell, Jean Alexandra, "Contextual Influences on Phonetic Categorization in Developmental Populations" (2013). *Master's Theses*. 394.
https://opencommons.uconn.edu/gs_theses/394

Contextual Influences on Phonetic Categorization in Developmental Populations

Jean Alexandra Campbell

B.A., University of Michigan, 2007

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts
At the
University of Connecticut
2013

APPROVAL PAGE

Master of Arts Thesis

Contextual Influences on Phonetic Categorization in Developmental Populations

Presented by

Jean Alexandra Campbell

Major Advisor ______________________________
Rachel M. Theodore, Ph.D.

Associate Advisor ______________________________
Bernard G. Grela, Ph.D.

Associate Advisor ______________________________
Tammie J. Spaulding, Ph.D.

University of Connecticut

2013

Acknowledgments

I would first like to acknowledge and thank Dr. Rachel Theodore for her encouragement and tireless support as my thesis advisor. Her zest for research, patience and support were instrumental in my choice to complete a Master's thesis and in carrying it through to completion. It is because of her support that I have gained the tools necessary to be a researcher.

I would like to acknowledge my research team, Alexis Giroux, Heather McSherry, and Vanessa Springer for their hard work on this project. Thank you for assisting me in all tasks, big and small! I'd also like to thank Kelly Casey and Marykate Bisaillon for providing technical support throughout this process. The lab would not run smoothly without you.

I would also like to acknowledge the efforts and support of Dr. Bernard Grela and Dr. Tammie Spaulding. Your comments and feedback have strengthened both my knowledge of the subject matter as well as this paper.

Finally, I would like to thank my family and friends for their encouragement throughout my graduate career. I sincerely appreciate all of your love and support. I would particularly like to thank my mother, Mary Jo Campbell, and my aunt, Claudia Oakes, for always taking the time to listen and for providing invaluable advice.

This research was supported by a University of Connecticut Research Foundation Faculty Large Grant and faculty start-up funds (College of Liberal Arts and Sciences, University of Connecticut) to Dr. Rachel M. Theodore.

Table of Contents

Abstract

A major goal of speech perception work is to describe how listeners extract individual consonants and vowels from the speech stream, and thus understand how listeners access the building blocks of larger linguistic units including words and longer utterances. A major challenge to this task is explaining how this is possible given the extreme variability present in the speech signal that prevents a one to one mapping of the acoustic signal to a particular speech sound. In healthy adults, findings have repeatedly demonstrated that the key to a healthy perceptual processing system is dynamically adjusting phonetic boundaries to accommodate contextual influences in speech production (Miller & Volaitis, 1989; Volaitis & Miller, 1992; Theodore, Miller, & Desteno, 2009). In other words, the process of efficiently extracting consonants and vowels from the speech signal requires a system that exhibits functional plasticity for systematic variation in the speech signal.

The current study seeks to determine if school-aged children demonstrate the same functional plasticity for systematic variation. We additionally hope to use the results of this study to motivate future work using the current paradigm with children with specific language impairment to provide more information about a possible processing-based underlying cause of the disorder. However, we face two challenges to addressing our questions. First, examining contextual influences in speech perception in children requires establishing a paradigm appropriate for this population. Second, a better understanding of the developmental time-course of contextual processing in typically developing children is necessary. The research presented in this thesis addresses these two challenges.

The purpose of Experiment 1 is to address the first challenge. It presents a study designed to test, with new adults, a paradigm that was designed for use with children. The goal was to develop a methodology to assess the contextual influences of speaking rate and place of articulation on phonemic boundaries. We tested adults to ensure that any modifications we made in the experiment design compared to earlier paradigms (e.g., Volaitis & Miller, 1992) still yielded the expected effect. Experiment 2 addresses the second challenge and presents a study that was designed to elicit baseline information for typically developing children.

Collectively, we found that older children demonstrated boundary flexibility similar to adults in both rate and place of articulation contexts. This demonstrates that these older children are taking contextual cues into account. For younger children, the results were less clear-cut and indicate that the paradigm may not be appropriate for young school-aged children.

The results of these two experiments added to our knowledge of language processing in three ways. First, the results indicate that the modified paradigm successfully measured categorical processing in healthy adults and typically developing children, ages 8-10. Second, the results also provided evidence in support of modifications to the paradigm, such as discrimination paradigms and imaging paradigms, to further assess the effects of context on the perceptual systems of younger children. Finally, the results provided evidence that the current paradigm will be appropriate for use with older children with SLI to help answer questions regarding the underlying cause of that disorder.

Introduction

Speech perception is generally considered to be the earliest stage of language comprehension (e.g., McClelland & Elman, 1986). It is at this stage that acoustic information is passed from the auditory nerve to higher-level structures for subsequent processing and interpretation. Functionally, research within the domain of speech perception aims to describe how listeners extract individual consonants and vowels from the speech stream, and thus understand how listeners access the building blocks of larger linguistic units including words and longer utterances. The fundamental challenge in this research domain is to describe how listeners achieve stable linguistic perception given that there is no one-to-one mapping between the acoustic signal and a given speech sound.

Research has demonstrated that adults with healthy processing systems are able to achieve stable perception in the face of rampant variability through a two-pronged process. The first is through the process of categorical perception or the ability of listeners to group acoustic signals into a category that is the matched to an individual speech sound (Lisker & Abramson, 1964; Eimas, Jusczyk, & Vigorito, 1971). The second part of the answer is that listeners are able to adjust the boundaries between these categories to account for changes in the acoustic signal (Miller & Volaitis, 1989; Volaitis & Miller, 1992; Theodore, Miller, & Desteno, 2009). These abilities are well documented in adults and category processing has also been found in infants as young as one month old (Eimas et al., 1971), however less is known about the speech processing abilities of school-aged children.

The purpose of this study is to develop a paradigm to test speech processing in the school-aged child population and to collect baseline data for the speech processing abilities of this population. We hope to then use both the paradigm and baseline information to assess speech-processing skills in school-aged children with specific language impairment (SLI). This is a language disorder that is found in approximately 7% of the United States population (Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997; Leonard, 1998) for which the underlying cause is not well understood. We hope that by adding to the knowledge base of the underlying cause of SLI we can help to improve diagnosis and treatment of this population.

To that end, we provide relevant background literature on speech perception in healthy populations and describe the current theoretical understanding of SLI. We conclude the Introduction by outlining empirical support for the prediction that some of the language deficits observed in SLI may reflect early speech processing deficits. Experiment 1 presents a study that assessed speech processing in adults using a paradigm that was designed for use with children in order to demonstrate the validity of our paradigm. Experiment 2 presents a study that examined typically developing children aged 5-10 years to determine baseline processing in that population. In the Discussion section, we outline the outcomes of Experiments 1 and 2 and the implications of these results on the experiment paradigm and our understanding of speech processing in typically developing children. Finally, we present future work that is motivated by the outcomes of these studies.

Speech perception

Before delving into an explanation of the current work, two aspects of language need to be considered: speech processing and speech production. At a basic level, speech processing is the system through which a listener receives acoustic information, recognizes that information as “speech,” decodes the acoustic information, assigns meaning to the information, and stores this information (Vance, Stackhouse, & Wells, 2005). Errors in speech processing can occur throughout this process if the acoustic information is not recognized, decoded incorrectly, or stored incorrectly. In contrast, speech production is the process by which a thought is conceived, syntactic rules are applied (e.g., part of speech), phonetic rules are applied, motor movements are initiated, and an acoustic form is produced and transmitted (Chomsky, 1965; Vance et al., 2005). Again, errors can occur at all levels of this process and will impact the signal transmission.

The current work considers the speech processing aspect of language, particularly the earliest stages that acoustic information must pass through before it can be decoded and understood by a listener. As mentioned earlier, a challenge of speech processing and for the listener is maintaining a stable and plastic system that is able to process the extreme variability present in incoming acoustic information. For example, consider the acoustic information used to specify just one class of sounds – stop consonants. The acoustic information produced for a given stop consonant varies widely due to a host of factors including speaking rate (Miller, Green, & Reeves, 1986; Nagao & de Jong, 2007), following phonetic context (Delattre, Liberman, & Cooper, 1955), place of articulation (Volaitis & Miller, 1992), and even who in particular is producing the stop consonant (Theodore, Miller, & Desteno, 2009). Thus the job for the listener in order to have stable linguistic perception is to take many different acoustic signals and map them to a single phonetic category.

Part of how listeners are able to accommodate variability for a given speech sound is their ability to “tune out” some of the variability and focus on a set of more salient characteristics. This ability gives rise to the concept of categorical perception. Though speakers actually produce individual phonemes with many slight acoustic variations, rather than mapping each individual variation to a separate phoneme, they map a range of variants onto a single phonemic category (Liberman, Harris, Kinney, & Lane, 1961). This allows all variations of /p/, for example, to be perceived and understood as that sound regardless of the talker, phonetic context, or speaking rate. This ability to perceive speech categorically develops early and research has found that infants as young as one month are able to approximate the categorical perception abilities of adults (Eimas et al., 1971).

One acoustic-phonetic property that has been examined with respect to categorical perception is voice-onset-time (VOT), which is a distinguishing feature of stop consonants (/p/, /b/, /t/, /d/, /k/, and /g/). These sounds are produced by completely occluding the vocal tract, and VOT is the amount of time that elapses between the release of the stop consonant and the onset of voicing that signals the beginning of the next sound. VOT is an important characteristic of these sounds because it distinguishes voiced stop consonants (/b/, /d/, /g/) from their voiceless counterparts (/p/, /t/, /k/). As shown in Figure 1, voiced stops typically have shorter VOTs than voiceless stops. Evidence for categorical perception comes from studies that have examined how listeners use VOT to recover voicing of word-initial stop consonants. In these experiments, listeners are presented with a range of VOTs and asked to identify the initial sound. Listeners hear VOTs that are typical of voiced stops (i.e., short VOTs) and voiceless stops (e.g., relatively long VOTs), but they also hear VOTs that are intermediate between these endpoints. Many studies have demonstrated that stop voicing perception is not linearly related to VOT. That is, in adults, the perception of voiceless stops does not get systematically higher in a linear relationship to VOT duration (e.g., Volaitis & Miller, 1992; Theodore et al., 2009). Rather, as illustrated in Figure 2, listeners perceive the VOT continuum categorically: a range of VOTs is identified as voiced, a different range of VOTs are identified as voiceless, and there is an abrupt discontinuity between the voiced and voiceless responses. Such categorical perception of acoustic-phonetic variation has been shown, in adults, for a host of phonetic properties that cue many different phonemic contrasts (Miller & Liberman, 1979; Volaitis & Miller, 1992; Schouten & Van Hessen, 1992; Miller, 2001; Minagawa-Kawai, Mori, & Sato, 2005).

Though categorical perception is a hallmark of efficient speech perception, many findings have demonstrated that the perception process is highly tuned to accommodate systematic acoustic-phonetic variation. That is, the precise boundaries between phonetic categories are not fixed in acoustic-phonetic space; rather, phonetic boundaries shift as a function of specific contexts. For example, consider the boundary specifying the voicing contrast in stop consonants. Research in adults has demonstrated that variability due to speaking rate and place of articulation can influence the VOT boundary. For example, as speaking rate slows, so does the location of the VOT boundary for stop consonants (Miller & Volaitis, 1989). Similarly, the voicing boundary for labial stops, which are produced at the front of the mouth, is located at a shorter VOT than the boundary for the posterior velar stops (Volaitis & Miller, 1992). These perceptual shifts in the voicing boundary are thought to occur in order to accommodate systematic variation in the speech signal; as speaking rate slows, VOTs produced for a given stop consonant increase (e.g., Theodore et al., 2009). As place of articulation moves from front to back in the vocal tract, VOTs increase in speech production such that VOTs for labial stops are typically shorter than VOTs for velar stops (e.g., Theodore et al., 2009). This makes for an efficient speech perception system that is able to quickly adjust phonetic boundaries to accommodate systematic changes in the acoustic signal so that comprehension of the incoming information continues even though acoustic variability is present.

While research has shown that adults are able to efficiently accommodate acoustic changes, this same pattern has not been studied in children. That is, we know that children are able to perceive speech categorically from a very young age, but we do not know at what age (Eimas et al., 1971) they begin to shift these boundaries to accommodate contextual variation. To sum, findings within the domain of adult and child speech perception have shown that listeners achieve stable perception by processing acoustic-phonetic variation categorically. Moreover, categorical processing in adults is flexible such that the precise boundary between phonetic categories is modified to accommodate contextual influences in speech production. The extent to which typically developing children exhibit such flexibility in categorical processing is currently unknown.

Specific language impairment

Specific language impairment (SLI) is a language disorder lag behind their peers in terms of language development in the absence of any known hearing, cognitive, obvious neurological deficits, or social-emotional deficits (Stark & Tallal, 1981; Leonard, 1998). Children with SLI have been observed to have impairments in phonology (Roberts, Rescorla, Giroux, & Stevens, 1998), morphology and syntax (Rice, Wexler, & Cleave, 1995; Leonard, 1999). Though the presentation of language deficits in children with SLI is well documented, the locus of the impairment is still a subject for debate. That is, even though the overt, observable atypical language skills are well documented, there is theoretical disagreement in terms of what aspect of the language processing is impaired and thus gives rise to these behaviors.

One theory holds that the locus of impairment for SLI is morpho-syntactic (Ullman & Gopnik, 1999; Rice et al.,, 1995). This theory proposes that the morphological and grammatical deficits witnessed in children with SLI are due to an incomplete understanding of the rules that govern language (van der Lely, 1996) or a subtle neurobiological deficit (Gopnik, 1997). In contrast, a competing theory places the locus of impairment at the level of speech processing (Joanisse & Seidenberg, 2003; Marinis, 2011). In other words, children with SLI do not correctly interpret or code incoming acoustic information. Further research is necessary in order to more definitively determine the root cause of this disorder, particularly with respect to these two competing theories. The goal of the current research is to provide a starting point for testing the hypothesis that the locus of language impairment in SLI stems from disruption early on in the processing stream. However, to do so we need to begin by establishing a paradigm to test this and by collecting baseline data from typically developing children to serve as our point of comparison.

Speech perception in children with SLI

The speech perception abilities of children with SLI have been studied in previous work and findings generally fall into two categories: abilities under optimal conditions and abilities under adverse conditions. Leonard and McGregor (1992) found that, when asked to perform a discrimination task, children with SLI performed similarly to typically developing peers when the sounds they were listening for had high phonetic substance (e.g., stressed syllables, longer duration, greater amplitude, higher frequency). However, for distinctions with less phonetic contrast, such as syllable final consonants and weak syllables, the children with SLI were less successful compared to typically developing children. Similarly, Coady, Kluender, and Evans (2005) found that children with SLI were able to discriminate between voiced and voiceless stop consonants when presented with clear examples of the given sounds. The children with SLI did not perform well when asked to differentiate between continua members that fell close to the VOT boundary. These findings may indicate that children with SLI do not have sufficiently stable category boundaries, and these boundaries can be disrupted in non-optimal conditions. Furthermore, Coady, Evans, Mainela-Arnold, and Kluender (2007) found that children with SLI performed speech discrimination tasks most successfully when the stimuli presented were real words and were produced by real speakers. Their task performance significantly decreased when stimuli were presented using synthesized speech and when non-word stimuli were used. This latter finding suggests that not only are children with SLI less stable in their categorical boundaries, but that they are also less flexible in accommodating changes in the acoustic signal (e.g., synthesized vs. natural speech). Overall, the body of research already conducted indicates that while children with SLI are able to perceive speech categorically, their ability to do so is much more fragile than that of their typically developing peers. If this is the case, then this research provides support for the speech perception theories of SLI impairment.

In summary, it is understood that VOT, a perceptual characteristic of stop consonants, is perceived categorically, VOT can be used to assess the speech perception abilities of children. The present study seeks to establish a paradigm and to collect baseline data to determine if school-aged children demonstrate the same categorical and plastic speech processing abilities that we see in healthy adults. Given that categorical perception is an early developing skill, we predict that the typically developing children will also demonstrate the same boundary flexibility as healthy adults during voiced and voiceless stop consonant discrimination tasks.

In the present work, experiment 1 seeks to establish a paradigm for testing our hypotheses regarding processing flexibility in an adult population to ensure that our framework (i.e. less repetitions) and stimuli (i.e. pictured response cues) elicit the expected results in adults. In Experiment 2 we will utilize the paradigm in Experiment 1 to establish a baseline for the boundary shifting abilities in typically developing children. The eventual goal is to use both the paradigm and baseline data to collect information about the same skills in the SLI population in order to add to the knowledge base regarding the underlying cause of that disorder. Experiments 1 and 2 could therefore be considered precursors to that later work.

Experiment 1

Experiment 1 was conducted with adult participants to test the research paradigm. The goal was to determine if changes made to previously utilized paradigms (e.g., Volaitis & Miller, 1992) would yield the same results. The most notable change to previously used paradigms concerns the number of trials presented; here we drastically lowered the number from 10 trials to five, in order to keep the time to complete the experiment more consistent with expectations for the developmental populations tested in Experiment 2. A second change concerns the fact that a new stimulus set was developed for this project; thus, we needed to ensure that the predicted results could be obtained with our particular set of stimuli.

Methods

Participants

Twenty-four adults ages 19-24 (mean = 21.12) participated in the experiment. Prior to beginning the experiment proper, the participant completed a short demographic form (see Appendix B). All listeners were either University of Connecticut students or residents living near the University of Connecticut. Five listeners were male and nineteen were female. Listeners were monolingual English speakers with no history of speech, language, or hearing disorders according to self-report, save for one adult who had a history of /r/ articulation disorder (for which she was successfully treated during childhood). All listeners passed a pure-tone hearing screen bilaterally at 20 dB HL for 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. Listeners were paid for their participation in the experiment.

An additional ten listeners were tested but not included in the analyses because they were not monolingual speakers (n = 8) or because of technical difficulties that resulted in the data from the experiment not being properly recorded during the experiment (n = 2).

Stimuli

The stimuli consisted of three VOT continua: a continuum from *goal* to *coal* at a slow speaking rate, a continuum from *bowl* to *pole* at a slow speaking rate, and a continuum from *goal* to *coal* at a fast speaking rate. Creation of the continua followed procedures outlined in Theodore and Miller (2010). To sum, the continua were based on natural productions of the voiced-initial endpoints *goal* and *bowl*. A female monolingual speaker of English was recorded producing many repetitions of these words (along with many fillers) and one repetition of each was selected such that word duration was approximately equivalent and the repetitions were of high acoustic quality (e.g., free from artifact). The selected *goal* and *bowl* tokens were equated for duration (*goal*, 478 ms; *bowl*, *479* ms), a cosine ramp was applied to the final 30 ms of each token in order to simulate the naturally-occurring decrease in amplitude at word-offset, and the two tokens were equated for root-mean-square amplitude.

A synthesized version of the selected *goal* and *bowl* tokens was created using LPC-based speech synthesis software (Analysis Synthesis Laboratory, Kay PENTAX) and this token served as the voiced-initial endpoint of the slow velar and slow labial continua, respectively. To create successive steps on each continuum, parameters of the LPC analysis were modified on a frame-by-frame basis (each frame corresponds to one vocal fold cycle) to replace the periodic source with a noise source and to scale peak amplitude by a factor of .22. After adjusting these parameters, a new token was synthesized based on the new parameters and the cycle was repeated. This procedure yielded, for each continuum, a series of tokens that incrementally increased in VOT in approximately 4 ms steps while maintaining constant word duration and filter characteristics of the original token. VOTs of the velar continuum ranged from 15-80 ms and VOTs of the labial continuum ranged from 10-75 ms; this offset with VOTs for the velar continuum slightly longer on a given step compared to the labial continuum is in accord with how these values pattern in natural speech (e.g., Theodore et al., 2009). Perceptually, the velar continuum ranged from *goal* to *coal* and the labial continuum ranged from *bowl* to *pole*. The fast velar continuum was created using the synthesized slow velar continuum. For each token on the slow velar continuum, 160 ms of vocal energy was deleted from the steady-state portions of the vowel and the final liquid at identical points for each token on the continuum. This procedure yielded a *goal* to *coal* continuum that had shorter token durations, and thus a perceptually faster speaking rate, but was in all other respects identical, to the slow velar continuum, most critically with respect to word-initial VOT. Tables 1 and 2 show VOT and word duration for each step of the three continua.

Procedure

Following a hearing screening, each listener was tested individually in a sound-attenuated booth. The listener was seated at a table that contained a computer monitor and a response box. Auditory stimuli were presented over headphones at a consistent volume as determine by computer settings. The experimenter delivered the instructions verbally. Directions were read from a prepared script to ensure that all participants heard exactly the same information.

All listeners participated in three blocks of trials, one for each test continuum. The order of the blocks was counter-balanced across listeners. Each block consisted of six randomizations of the particular continuum; the first block served as practice and was removed from further analysis. On each trial, listeners were asked to indicate whether the word began with a voiced or voiceless stop consonant by pressing an appropriately labeled button on the response box (Appendix A). Because this paradigm was being tested for use with children, the labels on the response box were pictures of the target word, and not orthography. Assignment of the voiced and voiceless labels was counter-balanced across participants to control for to dominant hand. Trials were presented with an ISI of 4000 ms, from button selection, and the next trial advanced if the listener failed to respond within 5000 ms of the onset of each trial. The experiment lasted approximately 30 minutes and the entire procedure, including hearing screen and completion of consent form and demographic forms lasted approximately 30-45 minutes.

Results

Identification responses were analyzed separately for each participant and for each continuum (i.e., fast velar, slow velar, slow labial) by calculating mean percent voiced responses for each step of the continuum across the five repetitions within each block. Two separate sets of analyses were performed, one to address the contextual influence of speaking rate and one to address the contextual influence of place of articulation.

Speaking rate

Figure 3 shows mean percent voiced responses across the 24 participants for the fast velar and slow velar continua. Consider first responses for each continuum. For both speaking rates, the expected categorical performance was observed such that a range of VOTs were consistently labeled as voiced, a different range of VOTs were consistent labeled as voiceless, and there was an abrupt discontinuity between the two ranges of VOTs. Moreover, inspection of the figure suggests that the boundary for the slow speaking rate continuum is indeed located at a longer VOT compared to the boundary for the fast speaking rate continuum; that is, the entire function is displaced towards longer VOTs. This is the predicted influence of speaking rate on perception of the voicing contrast.

To measure the statistical significance of this effect, we calculated the boundary for each participant, for each continuum, using standard conventions as outlined in Miller and Volaitis (1992). First, an ogive function was fit to the identification responses using probit analyses. Essentially, this process fits the identification data to a cumulative normal distribution using a regression model. Second, the mean of this distribution, that is the VOT corresponding to 50%, was used as a metric for the category boundary. The ogive function was a good fit to the raw data for all participants for both the slow and fast continua, with r values ranging from 0.93 to 1.00.

Figure 4 shows the mean boundary for the slow and fast continua across participants. A paired t-test confirmed that the mean boundary was located at a longer VOT for the slow compared to the fast continuum, indicating that listeners adjusted the phonetic boundary to accommodate the contextual influence of speaking rate [$t(23) = 6.07$, $p < .001$].

Place of articulation

Figure 5 shows mean percent voiced responses across the 24 participants for the velar (*goal* to *coal*) and the labial continua (*bowl* to *pole*). Each continuum exhibited the expected pattern of categorical responses and the ogive function was a good fit to the raw data for all participants for both the velar and labial continua with r values ranging from 0.96 to 1.00. Moreover, the continuum is displaced such that the identification function is shifted toward longer VOTs for the velar compared to the labial continuum. To examine the statistical significance of this pattern, boundaries were calculated for each participant, for each continuum, following the curve-fitting procedure outlined above. Figure 6 shows the mean boundaries across participants and a paired t-test confirmed that the velar voicing boundary was indeed located at a longer VOT compared to the labial voicing boundary [$t(23) = 24.93$, $p < .0001$]. This pattern of results indicates, as predicted, that listeners dynamically adjusted the voicing boundary to accommodate the contextual influence of place of articulation.

Conclusions

As stated in the Introduction, the goal of Experiment 1 was to test the proposed paradigm to ensure that changes made to the stimuli and procedures still yielded the predicted boundary shifts in healthy adults. The results indicate that the expected boundary shift was present for both contextual influences of speaking rate and place of articulation.

Though the desired effect was present in both contexts, it should be noted that numerically, the difference between the mean VOT for the different continua was greater for the place effect than for the rate effect. This does not necessarily mean that listeners are more sensitive to differences in place context but could instead be the result of the adjustments we made to simulate changes in speaking rate were not equal to the magnitude of change for our place of articulation manipulation. In other words, we do not know that the metric for speaking rate is the same as it is for place of articulation, and thus it is not appropriate to compare the magnitude of the rate effect to the magnitude of the place effect. That is, the two contextual influences are qualitatively different. To underscore this point, it is possible that if the difference in word duration of the “fast” and “slow” continua was increased substantially, then the magnitude of the displacement for the rate continua might be more on par with the displacement observed for the current place of articulation continua.

In summary, the results of Experiment 1 indicate that the paradigm that we have established yielded the expected boundary shifts in the healthy adult population. This indicates that the changes we made to the stimuli and procedure did not impact boundary flexibility and suggests that the paradigm should yield the same results in a child population if they do in fact demonstrate the same boundary plasticity as adults. What follows in Experiment 2 is an experiment testing just this hypothesis in typically developing children.

Experiment 2

As stated in the Introduction, while it has been established that adults demonstrate great functional plasticity in terms of their speech processing, this same pattern has not been widely tested in children. We do know that infants begin to show categorical perception of phonetic variation from a very young age (Eimas, 1971), which raises the possibility that they also develop the same boundary flexibility as adults at a young age. Having established an appropriate paradigm to test the effect of rate and place of articulation contexts on the voicing boundary of stop consonants, Experiment 2 uses this same paradigm with typically developing children to establish feasibility of the paradigm with a child population and to collect baseline data for voicing boundary shifts in response to contextual changes in typically developing children.

Methods

Participants

Twenty children were recruited from the University of Connecticut community. Due to the pilot nature of this study, all children who met the age and hearing screen criteria were eligible for participation. All child participants were between the ages of 5-10 years and were monolingual native English speakers. Twelve of the children who participated were male and eight were female. All children participated in a pure-tone hearing screen and all but one child (who had a previously diagnosed unilateral hearing loss) passed bilaterally; this child did pass the screening at all frequencies unilaterally, and was thus included in the study. Parents completed a demographic form for each child participant (Appendix B). Two children had previous diagnoses of speech and language difficulties (both Childhood Apraxia of Speech); no other child had a history of speech, language or cognitive deficits, according to parent report. It should be noted that three of the children declined to continue the experiment; therefore, we do not have complete data from them. For three different children, there were technical difficulties during data collection and consequently their data sets were incomplete. As described below, all children participated in a category identification test and completed a testing battery that assessed language, articulation, phonological, and nonverbal cognitive abilities. Children received trinkets as a reward for participating and families were provided with monetary compensation to offset transportation costs.

Stimuli

The stimuli described in Experiment 1 were also used in Experiment 2.

Procedure

The procedures outlined in Experiment 1 were followed with three exceptions. First, auditory stimuli were presented at a consistent level as determined by computer settings, via a single speaker placed approximately 24 inches in front of the child instead of via headphones due to concerns that young children could not tolerate wearing headphones for the entire length of the experiment. Second, prior to the experiment proper, the researcher checked to ensure that the participant understood the target vocabulary. This process consisted of pointing the examiner pointing to each picture, labeling it, providing a short description of the word (e.g., “This is a *goal*. It’s where you kick a ball to score in soccer.”), and then asking the child to point to the pictures as the examiner named them. All children were able to point to all pictures following the examiner request and we took this as evidence that the children could match the auditory stimuli to pictures cues sufficient to participate in the experiment.

Finally, each child participated in the speech, language, and cognition assessment battery following the conclusion of the category identification test. The tests were administered following the identification test to ensure that the children were most engaged and alter during the actual experiment task. The *Goldman-Fristoe Test of Articulation – Second Edition* (GFTA-2*;* Goldman & Fristoe, 2000) was administered as a measure of articulation. *The Kahn-Lewis Phonological Analysis – Second Edition* (KLPA-2; Khan & Lewis*, 2002*) was applied to the *GFTA-2* to assess the presence of phonological processes. *The Clinical Evaluation of Language Fundamental – 4 Screening Test* (CELF-4 Screening Test; Semel, Wiig, & Secord, 2004) was used to screen for expressive and receptive language impairment. Finally, *Raven's Coloured Progressive Matrices* (Raven's; Raven. Raven, & Court, 1998) was administered as an assessment of nonverbal cognitive skills. All assessments were administered in the order presented above and all but two assessment batteries were video recorded. Twenty-five percent of the assessments were reviewed and scored by a second undergraduate or graduate experimenter; reliability between the two experimenters was high (89%). The experiment lasted approximately 30 minutes and the assessment battery, hearing screen, and form completion lasted 30-45 minutes, though this time varied widely across the children.

Results

As in Experiment 1, identification responses were analyzed separately for each participant and for each continuum (i.e., fast velar, slow velar, slow labial) by calculating mean percent voiced responses for each step of the continuum. Each set of identification data was submitted to curve fitting in order to quantify the voicing boundary for each continuum. In contrast to the adult data presented in Experiment 1, there was wide variability in the goodness of fit of the ogive function to the identification data across the children. Figure 7 shows representative curve fits for two children, one that highlights an extremely poor fit and one that highlights a good fit. Because our method of calculating the voicing boundary hinges on the ogive function providing a veridical representation of identification responses, children for whom the ogive was a poor fit to their identification responses ($r < .70$) were excluded from the primary analyses. A poor ogive fit indicates that the children's responses were not categorical. Since, as mentioned above, our boundary calculations rely in the mean provided by the ogive, these children had to be excluded because we could not say with confidence that they moved their boundaries if we were not sure they were stable We address this issue in greater detail below. As in Experiment 1, two separate sets of analyses were performed, one to address the contextual influence of speaking rate and one to address the contextual influence of place of articulation.

Speaking rate

Of the 20 children who participated in this experiment, only 11 were included in the final analyses (see Appendix C). Six children were excluded because they did not complete all test trials and three were excluded because the ogive was a poor fit to their identification responses. Figure 8 shows mean percent voiced responses for the 11 children included in this analysis for the fast and slow speaking rate continua. Three observations should be noted: (1) the identification functions for both speaking rates are less categorical than was observed in Experiment 1, particularly in that the endpoint continua do not reach ceiling and floor responses; (2) variability as shown by standard error of the mean is greater for the child data compared to the adult data presented in Experiment 1; and (3) the two functions are displaced in the appropriate direction, with the slow speaking rate shifted toward longer VOTs compared to the fast speaking rate.

As in Experiment 1, a boundary value for each child, for each continuum, was obtained using probit analyses. Figure 9 shows the mean boundaries across the children for the fast and slow continua. Results of a paired t-test indicated that there was no significant difference between boundaries of the fast and slow speaking rate continua, indicating that these children did not shift the voicing boundary in accord with the contextual influence of speaking rate [$t(10) = -1.55$, $p = .151$].

One last t-test was performed on the speaking rate boundaries. Recall that a criterion of $r > .70$ was set for goodness-of-fit between the observed responses and those fitted with the probit analyses. As described in Experiment 1, this criterion is well below the correlations for the fitted curves observed for the adults.

In order to examine how the contextual influence of speaking rate might affect boundary for those children who have identification functions that were as “categorical” as the adults, a paired-test was used to compare the boundary of the fast and slow speaking rates for the five children who had *r* values within the range of adults. Despite the low-powered analysis, the results showed that the boundary for the fast continuum was located at a significantly shorter VOT compared to the slow continuum [t(4) = -2.76, p = .050].

Place of articulation

Of the 20 children who participated in this experiment, only 13 were included in the final analyses (See Appendix C). Four children were excluded because they did not complete all test trials and three were excluded because the ogive was a poor fit to their identification responses ($r < .70$). Figure 11 shows mean percent voiced responses for the 13 children included in this analysis for the labial and velar continua. As for the speaking rate data presented above, both the labial and velar identification functions show increased variability relative to the corresponding adult data. However, as opposed to the speaking rate functions, the labial and velar identification functions show substantial displacement, with the velar continuum shifted toward longer VOTs compared to the labial continuum. Mean boundaries for the labial and velar continua as determined from the probit analyses are shown in Figure 10 and were submitted to a paired t-test. The results illustrate that the boundary for the velar continuum was located at a significantly longer VOT compared to the labial continuum [t(12) = -8.35, p < .001]. This finding indicates that the children processed VOT with respect to place of articulation, and thus demonstrated sensitivity to this contextual influence for perception of the voicing contrast.

Conclusions

The purpose of Experiment 2 was twofold: The first aim was to assess the feasibility of the current paradigm for use with children in the current population and the second was to gather baseline data regarding the VOT boundary flexibility in typically developing children. While we gained valuable information with respect to our two goals, the wide variability in performance of the children requires some discussion and interpretation.

We will begin by considering the second goal: gathering baseline data for the typically developing population. We observed a lot of variability in the degree to which identification functions – for a given continuum – fit a standard categorical response pattern. Only 11 of the children performed in this fashion for the speaking rate continua and only 13 of the children performed in this fashion for the place of articulation continua. Because examining contextual influences in speech perception requires comparing the boundary between two contextual conditions, a prerequisite for examining contextual processing is the ability to determine a stable acoustic-phonetic boundary. That is, a failure to observe contextually-driven boundary shifts can only be interpreted to the degree that the boundaries in question are validly measured. With respect to this, the results demonstrated that children who did in fact respond categorically did show evidence of flexibility for the place of articulation context. This indicates that children in the current population do show evidence of the ability to shift their perceptual boundaries with regard to changes in place of articulation context. The results, however, should to be interpreted with a fair amount of caution given the small sample.

With regard to the feasibility of the use of current paradigm in the child population, the current experiment yielded mixed results. Informal analyses of the child data presented in the Appendix C suggest that for older children between the ages of 8-10, this paradigm is a valid way to measure categorical perception. Children in this age group showed categorical responses, using *r* as an indicant of goodness-of-fit, between observed and fitted identification functions. However, the children who were excluded from the results because they were not providing stable identification functions were at the younger ages of our test group. As previous research has proven that children should perceive speech categorically by this age (Eimas, 1971), we do not take these results to indicate a lack of categorical ability in these children. Rather, the results suggest that for the younger children in our study, the paradigm in not effectively assessing categorical processing. This in turn means that the issue here is one of validity and that we have not measured what we hoped to for these children.

In summary, we made some progress toward our two stated goals. With respect to the first, we determined that the present paradigm is appropriate for some children, namely older children. This suggests that the current paradigm may require adjustment to be used with a broader child population. In terms of the second question, we were able to determine that there is evidence to suggest that children in the present population do demonstrate some flexibility in their perceptual boundaries, indicating that such functional plasticity of phonetic boundaries is acquired early in the developmental trajectory.

Discussion

The purpose of this study was two-fold. The first we sought to develop a paradigm of assessing boundary flexibility in typically developing children. Secondly we hoped to use this paradigm to begin collecting data from typically developing children. The study is motivated by the paucity of research regarding functional plasticity of phonetic representations in children. Experiment 1 represents our attempt to establish a paradigm that elicits the desired flexibility patterns in healthy adults. Experiment 2 represents our attempt to assess the feasibility of the paradigm for use with a child population. Experiment 2 is also our attempt to begin to provide data regarding the plasticity of speech processing in typically developing children.

Research in the domain of speech perception has provided a great deal of evidence to motivate the current research. We know that listeners process the speech stream categorically from a very young age in order to map the acoustic information to individual phonemes (Eimas et al., 1971). This allows listeners to achieve stable perception despite variability in the speech signal. Furthermore, we know that adult listeners are quite flexible in the location of these boundaries and can adjust them to account for context changes (e.g., speaker, rate, place, phonetic context) in the incoming speech signal (Miller et al., 1986; Volaitis & Miller, 1992; Miller, 2001; Theodore et al., 2009). Given that infants as young as one month old (Eimas et al., 1971) have been noted to perceive speech categorically, we predicted that typically developing children would demonstrate boundary plasticity similar to the adult participants.

The present study provided partial answers to our proposed questions. Our paradigm for assessing category flexibility as a function of speaking rate and place of articulation context was successful for adults. Additionally, it appears that the paradigm is appropriate for some children, particularly those ages 8-10 years. A larger sample is necessary to confirm these results, but our initial findings are promising. Additionally, we were able to preliminarily conclude that, particularly for place of articulation, some typically developing children have acquired the ability to flexibly shift categorical boundaries. Again, a larger sample size will be essential in confirming this, but the present study does suggest that further research in this area is warranted. In summary, while additional participants are required to confirm the tentative initial findings discussed here, the results provide preliminary evidence suggesting that the proposed paradigm is appropriate for some children.

Future Work

Paradigm

Due to the fact that the paradigm was not appropriate for all children, the present study also provides some insight into ways the current paradigm could be monitored to examine the same concept more successfully in the younger population. There are two adjustments that could be made to the paradigm that might be appropriate. The first would be to present the experiment as a behavioral discrimination task in which children would be presented with two stimuli from different points on the continuum and asked to determine if the presented words are the “same” or “different.” These types of tasks have been used with relative success with children and therefore might be appropriate for the population in question (Coady et al., 2005; Coady et al., 2007). Attempts are made to reduce memory demands during these tasks by using small ISI durations (Coady et al., 2005; Coady et al., 2007). These tasks do however have pitfalls of their own. In their study using discrimination tasks, Coady and colleagues (2005) did not include children below the age of 6 years, 11 months, and still commented that some variability may be the result of inattention. Furthermore, discrimination paradigms require the use of the abstract concepts of “same” and “different,” which may be difficult for young children to comprehend.

Another possibility that would seek to circumvent some of these issues would be to use electrophysiological discrimination tasks. In these tasks, neural responses in the brain are measured to determine if the brain is detecting perception of a difference in a chain of stimuli (Kraus, McGee, Carrell, King, Tremblay, & Nicol, 1995; Sharma & Dorman, 1999; Uwer, Albrecht, & von Suchodoletz, 2002). This allows us to determine whether the brain is interpreting input differences, even if the listener is not consciously aware of the difference or is unable to demonstrate an awareness of the difference (e.g., due to age, task demands). The passive nature of the task offers significant benefits for young children, who will likely be unreliable in tasks requiring a consistent response. This type of experiment design has been used successfully for speech perception tasks in very young children in the past (Kraus et al., 1995) and in speech processing tasks with children with SLI (Uwer et al., 2002). Extensions of the current work in behavioral discrimination and ERP paradigms are currently in progress.

SLI Population

As mentioned in the introduction, a motivation for our study was to establish a paradigm for assessing speech processing in children with SLI. Research has established that children with SLI demonstrate deficits in all facets of language: phonology, morphology, semantics and syntax, though trouble with morpho-syntax is considered to be the hallmark characteristics of this disorder (Leonard, 1998). Though the previously mentioned characteristics of the disorder are generally considered to be common among children with SLI, the SLI population is actually quite diverse (Conti-Ramsden & Botting, 1999). The heterogeneity of the population, along with the fact that the particular constellations of strengths and deficits can change throughout a person's lifetime, makes differential diagnosis of SLI a challenge. Most frequently, an SLI diagnosis is made through a process of ruling out other possible causes of the deficits.

Further complicating the diagnosis process is the fact that researchers and speech-language pathologists in the field are not in agreement regarding the etiology of the disorder (Joanisse & Seidenberg, 1998). Research does suggest that the disorder likely includes a genetic component, as a person is more likely to be diagnosed with SLI if he or she has a family member with the disorder (Bishop, 2006). Less clear, however, is the exact locus or root of the impairment. Two prevailing schools of thought regarding the locus of impairment have a body of evidence supporting them and were touched upon in the introduction. They are discussed in more detail below.

In the first theory of SLI with evidence, the locus of the disorder is considered to be one of speech production and at the syntactic or grammatical level (Rice et al., 1995; Ullman & Gopnik, 1999; van der Lely & Stollwerk, 1997). Syntactic-based theories build on Chomsky's theory of an underlying universal grammar present in all children at birth. Chomsky's theory of universal grammar holds that children have an innate understanding of the structure of language (Chomsky, 1965). Proponents of syntactic theory of SLI deficit argue that children with SLI demonstrate morpho-syntactic errors that are inconsistent with processing or articulation deficits. For example, Ullman and Gopnik (1999) found that participants with SLI struggled to apply regular past tense and third person singular (i.e., *–ed* and *–s*) to nonsense verbs but demonstrated stronger ability to utilize irregular past tense for nonsense verbs. This suggests that the underlying rules for regular past tense and plural are not fixed in the SLI population.

One issue with regard to morpho-syntax-based theories is that they do not adequately account for the fact that children with SLI do produce many correct syntactic structures but are either inconsistent in their productions or struggle with certain structures (Joanisse & Seidenberg, 2003). This suggests that a model based on missing underlying structures cannot completely explain the deficits witnessed in the SLI population.

Another theory places the locus of the impairment at the perceptual level of communication. The speech processing deficit theory attempts to account for the impaired linguistic behavior observed in people with SLI as well as aspects of their communication abilities that are not adequately addressed by other theories. Several sub-theories have been proposed for the specific type and location of breakdown in processing. The one we will consider here points to deficits in phonological working memory. When phonologic information enters the auditory system it must be held until it can be analyzed and stored in long-term memory. Researchers who subscribe to a phonological memory deficit argue that for people with SLI, phonological information is either stored incorrectly during encoding, is degraded before it is filed into long-term storage, or is completely lost before it reaches long-term storage (Gathercole & Baddeley, 1990).

Joanisse and Seidenberg (2003) provide support for a phonological memory breakdown and a framework for understanding how deficits in phonological working memory can also lead to deficits at the level of syntax production. They were able to create a connectionist model for pronoun comprehension that mimicked impaired phonological working memory by introducing phonemic variability during input. They then “trained the system” and tested it through comprehension and production simulations. They were able to generate syntactic errors consistent with those produced in the SLI population by simulating an impaired phonological working memory. They determined that the model provides evidence for how perception and phonological representations can impact syntactic production. In summary, research in support of a processing-based deficit suggests that the language processing capabilities of children with SLI warrant further analysis.

We believe that the current paradigm would be appropriate for use to assess boundary flexibility in older children with SLI. Assessing boundary flexibility would be one avenue for adding to the knowledge base regarding SLI locus of impairment. We predict that children with SLI will not demonstrate the same boundary flexibility as their typically developing peers. Motivating for this come from studies mentioned in the introduction (Coady et al., 2005; Coady et al., 2007) which demonstrate that children with SLI have less stable category boundaries and are less flexible in adjusting them to accommodate less optimal acoustic conditions. If the children do demonstrate deficits in boundary flexibility tasks, this provides evidence in support of a processing deficit which will help to inform diagnosis and treatment of this disorder.

Clinical Implications

In the short-term the current work provides evidence that by approximately age eight, typically developing are showing adult-like abilities to adjust perceptual boundaries in response to contextual changes in the acoustic signal. Children in this age range who are not able to do this are therefore deficient in their processing abilities. Clinically, if a child is older than eight-years-old, but demonstrates a deficiency in adjusting categorical boundaries to accommodate context, the clinician needs to include activities that ensure sufficient understanding of incoming information. After all, if information is not learned correctly, it is unlikely that it will be used correctly.

In the long-term the results of this research will help to drive diagnosis and treatment of children with SLI. For example, if the locus of impairment were at the level of morpho-syntactic knowledge, then therapy should target strengthening the underlying structural framework by teaching the missing or weak syntactic structures (Yoder, Molfese, & Gardner, 2011). Alternatively, if the disorder is processing based, then treatment should target initial comprehension of incoming acoustic information with the expectation that this would in turn facilitate production of the correct morpho-syntax. We believe that the current paradigm is nearly ready for use with the 8-10 year old SLI population and the results of that research will help to answer questions regarding the underlying cause of SLI.

References

Bishop, D. M. V. (2006). What causes specific language impairment in children? *Current Directions in Psychological Science, 15*, 217-221.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The M.I.T Press.

Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2007). Children with specific language impairments perceive speech most categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research, 50,* 41-57.

Coady, J. A., Kluendrer, K. R., & Evans, J. L. (2005). Categorical perception of speech by children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 48*, 944-959.

Conti-Ramsden, G. & Botting, N. (1999). Classification of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1195-1204.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27*, 769-773.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303-306.

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

Goldman, R. (2000). *Goldman Fristoe test of articulation*. Circle Pines, MN: American Guidance Service.

Gopnik, M. (1997). Language deficits and genetic factors. *Trends in Cognitive Sciences, 1*, 5-9.

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

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language, 86*, 40-56.

Khan, M. M., & Lewis, N. (2002). *Khan-Lewis phonological analysis*. Circle Pines, MN: American Guidance Service.

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7*, 25-32.

Leonard, L.B. (1998). *Children with specific language impairment*. Cambridge, MA: The MIT Press.

Leonard, L. B. (1999). Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 678-689.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 741-753.

Leonard, L. B. & McGregor, K. K. (1992). Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech & Hearing Research*, *35*, 1076-1085.

Leonard, L. B., Miller, C. A., Grela, B. G., Holland, A. L., Gerber, E., & Petucci, M. (2000). Production operations contribute to the grammatical morpheme limitations of children with specific language impairment. *Journal of Memory and Language, 43,* 362-378.

Liberman, A. M., Harris K. S., Kinney, J., & Lane H. (1961). The discrimination of relative onset-time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology, 61*, 379-388.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustic measurements. *Word, 20, 384-422.*

Marinis, T. (2011). On the nature and cause of specific language impairment: A view from sentence processing and infant research. *Lingua*, *121*, 463-475.

McClelland, J. L., & Ellman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1-86.

Miller, J. (2001). Mapping from acoustic signal to phonetic category: Internal category structure, context effects and speeded categorisation. *Langauge and Cognitive Processes, 16*, 683-690.

Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica*, *43*, 106–115.

Miller, J. L., & Liberman, A. M. (1979) Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, *25*, 457-465.

Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, *46*, 505-512.

Minagawa-Kawai, Y., Mori, K., & Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes. *Journal of Cognitive Neuroscience, 17,* 1376-1385.

Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research*, *38*, 187-199.

Nagao, K., & de Jong, K. (2007). Perceptual rate normalization in naturally produced rate-varied speech. *Journal of the Acoustical Society of America*, *121*, 2882–2898.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Coloured Progressive Matrices*. Bloomington, MN: Pearson.

Rice, M.L., Wexler, K., & Cleave, P.L. Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research,* 38*,* 850-863.

Roberts, J., Rescorla, L., Giroux, J., & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E): Outcome at age 3. *Journal of Speech, Language & Hearing Research, 41*, 374-384.

Schouten, M.E.H. & van Hessen, A.J. (1992). Modeling phoneme perception. I: Categorical Perception. *Journal of the Acoustical Society of America*, *92*, 1841-

1855.

Semel, E., Wiig, E. H., & Secord, W.A. (2004). *Clinical Evaluation of Language Fundamentals, Screening Test, Fourth Edition*. San Antonio, TX: Pearson.

Sharma, A., & Dorman, M. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America*, 106, 1078-1083.

Stark, R.E. & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders,* 46, 114-122.

Theodore, R., Miller, J., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: contextual influences. *Journal of the Acoustical Society of America*, *125,* 3974-3982.

Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *Journal of the Acoustical Society of America, 128,* 2090-2099.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language & Hearing Research, 40*,1245–60.

Ullman, M. T. & Gopnik, M. (1999). Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics, 20*, 51-117.

Uwer, R., Albrecht, R., & von Suchodoletz, W. (2002). Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine & Child Neurology, 44*, 527-532.

Vance, M., Stackhouse, J., & Wells, B. (2005). Speech-production skills in children aged 3-7 years. *International Journal of Language & Communication Disorders, 40*, 29-48.

Van der Lely, H.K.J. (1996). Specifically language impaired and normally developing children: verbal passive vs. adjectival passive sentence interpretation. *Lingua*, 98, 243-272.

Van der Lely, H.K.J., & Stollwerk, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, *62*, 245-290.

Volaitis, L. E. & Miller, J. L, (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America*, *92*, 723-735.

Yoder, P. J., Molfese, D., & Gardner, E.. (2011). Initial mean length of utterance predicts the relative efficacy of two grammatical treatments in preschoolers with specific language impairment. *Journal of Speech, Language and Hearing Research, 54*, 1170-1181.

Figures

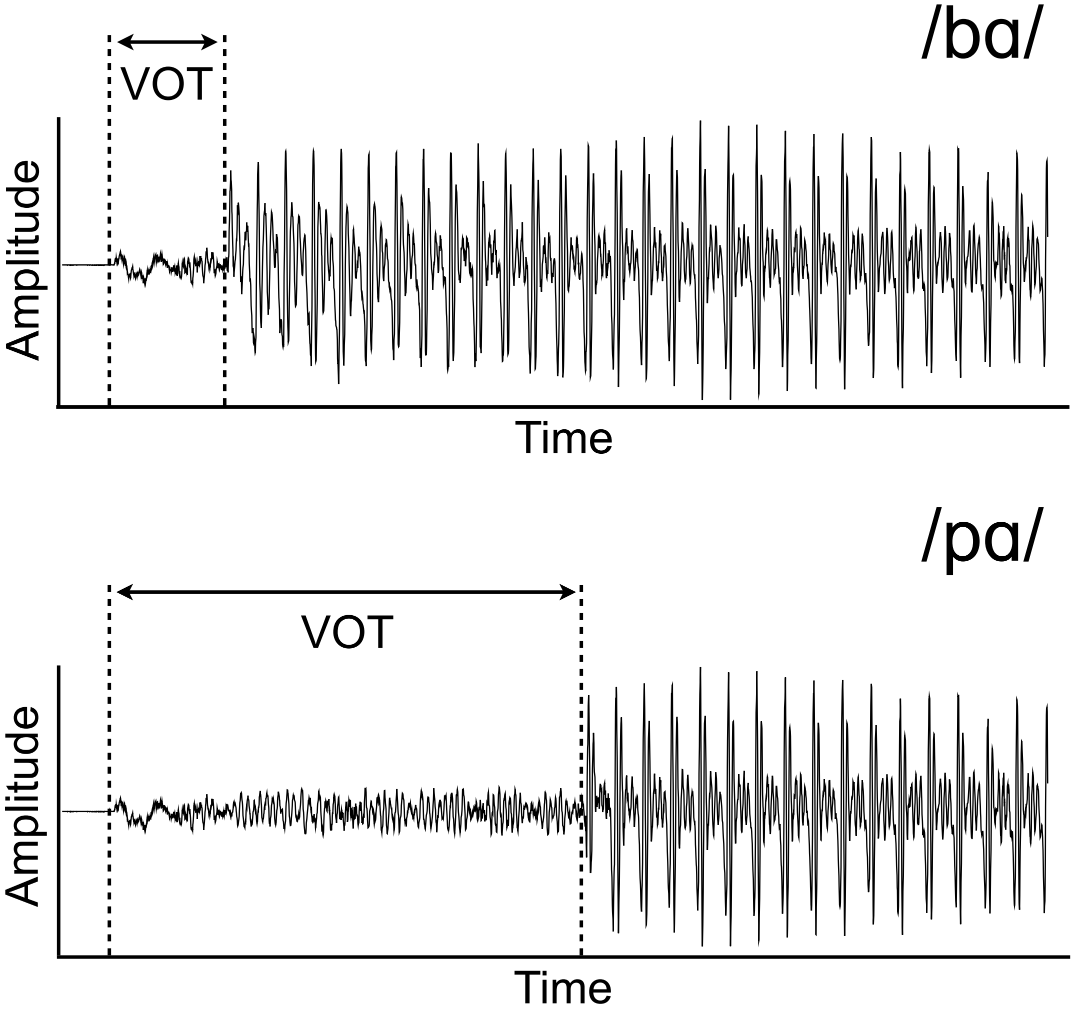

Figure 1. Representative waveforms showing voice-onset-time for a voiced stop (top panel) and a voiceless stop (bottom panel).

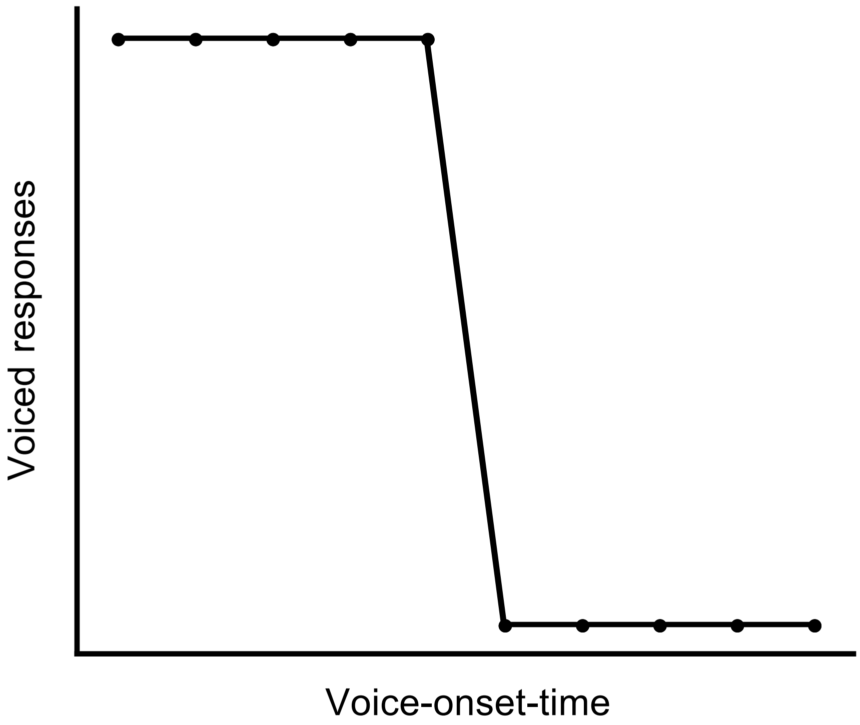

Figure 2. Representative identification function illustrating categorical perception of voice-onset-time.

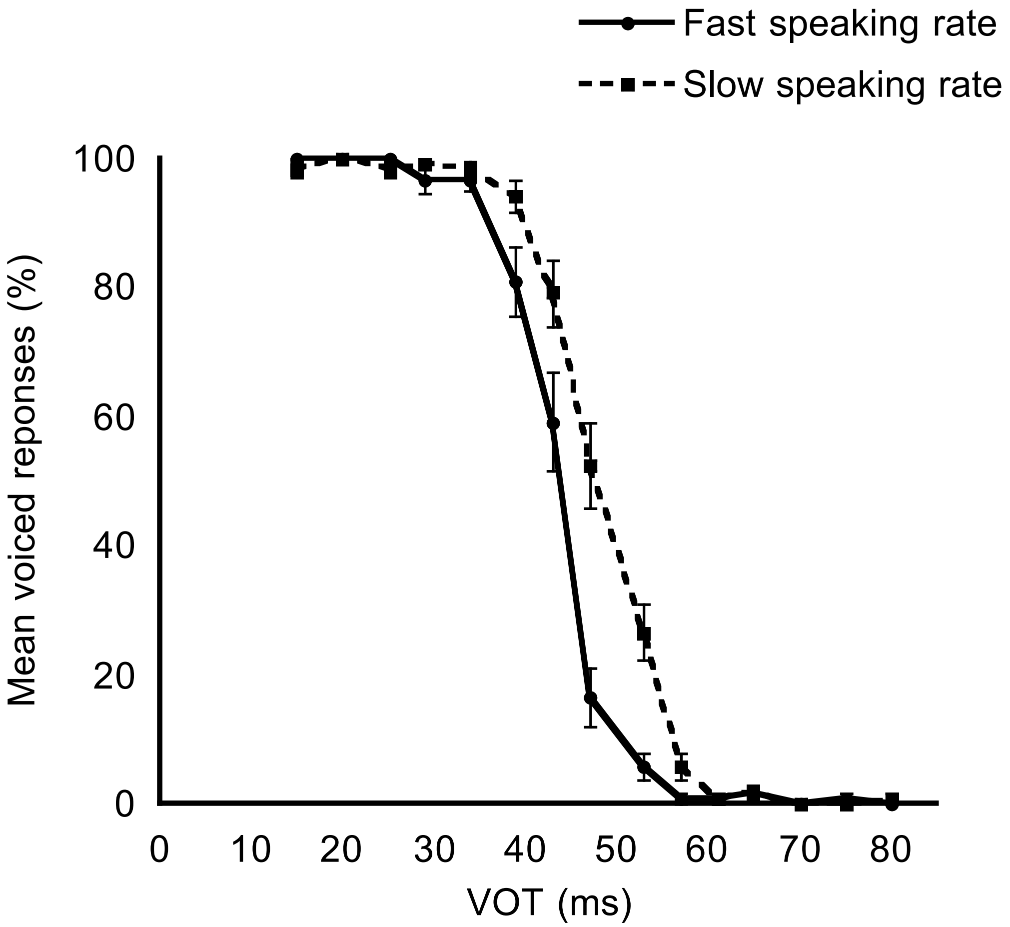

Figure 3. Mean percent voiced responses as a function of voice-onset-time for the fast and slow speaking rates tested in Experiment 1. Error bars indicate standard error of the mean.

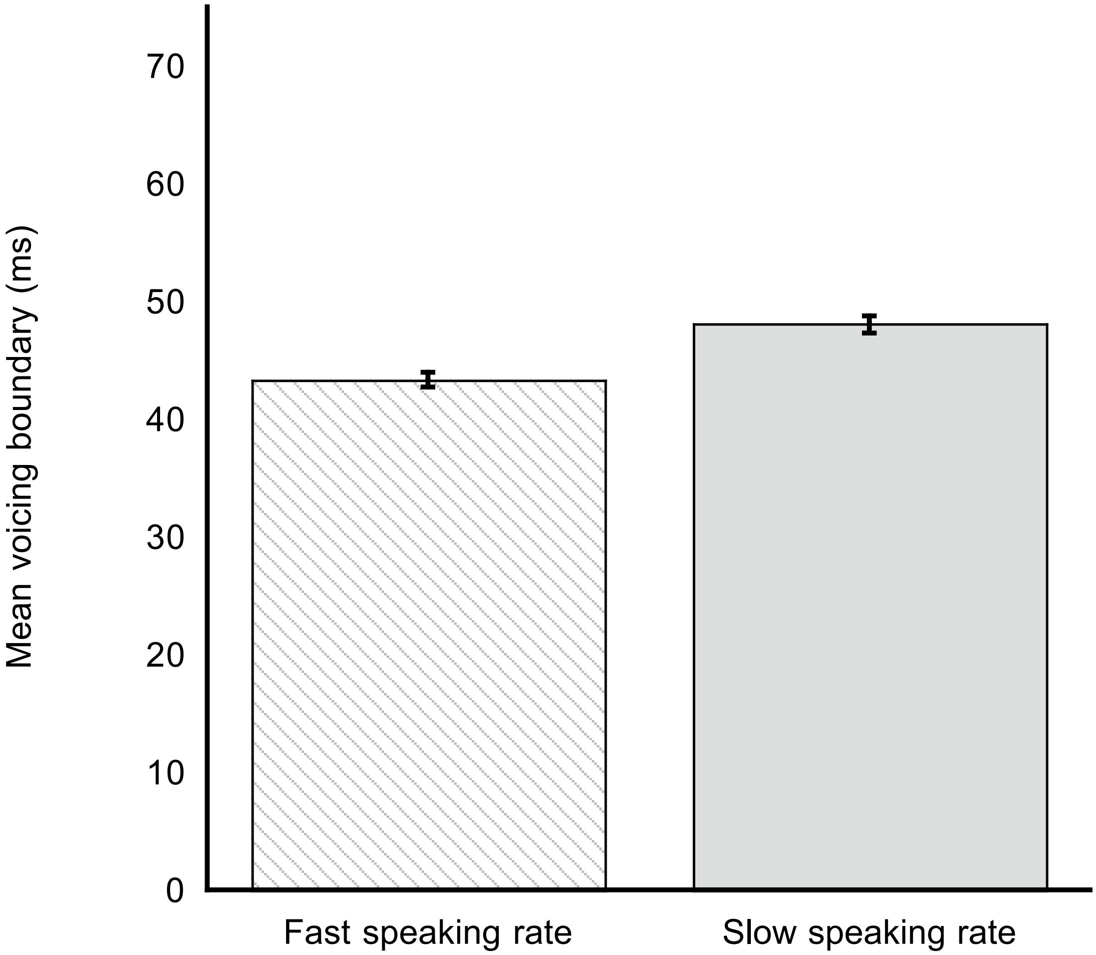

Figure 4. Mean voicing boundary for the fast and slow speaking rates tested in Experiment 1. Error bars indicate standard error of the mean.

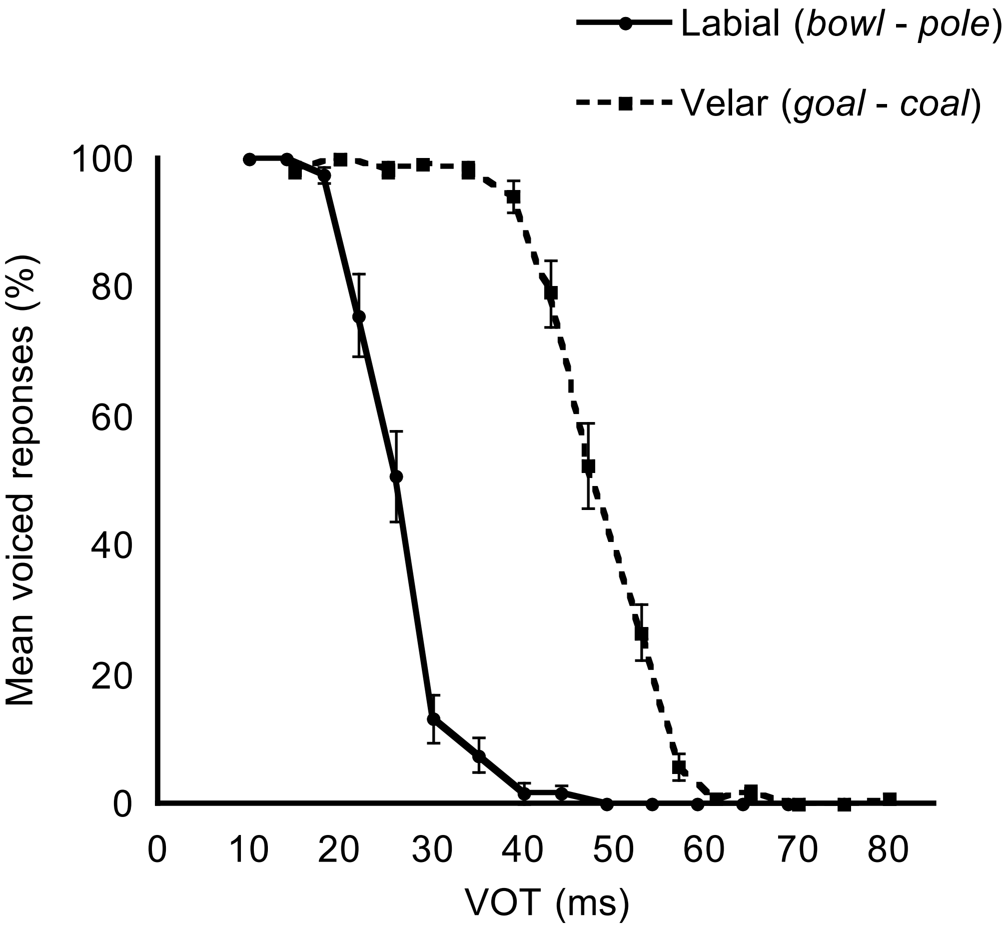

Figure 5. Mean percent voiced responses as a function of voice-onset-time for the labial and velar continua tested in Experiment 1. Error bars indicate standard error of the mean.

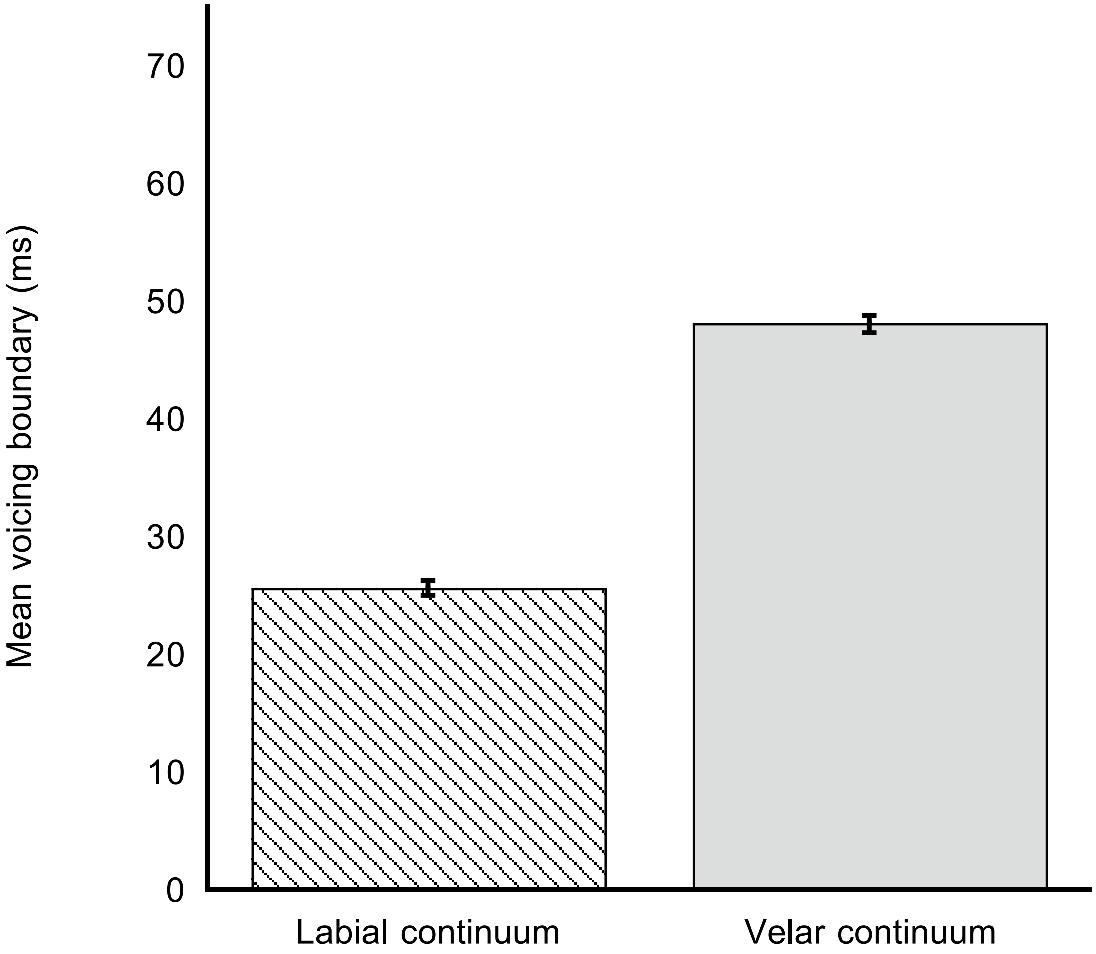

Figure 6. Mean voicing boundary for the labial and velar continua tested in Experiment 1. Error bars indicate standard error of the mean.

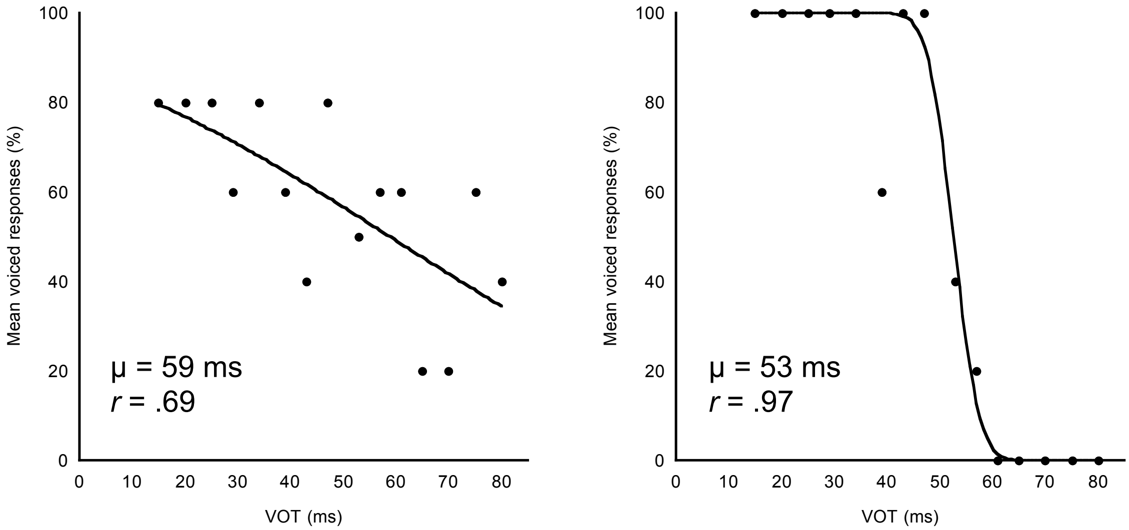

Figure 7. Representative identification functions illustrating a poor fit to the ogive (left panel) and a good fit to the ogive (right panel). Individual data points are shown in filled circles and the line in each panel shows the fitted curve.

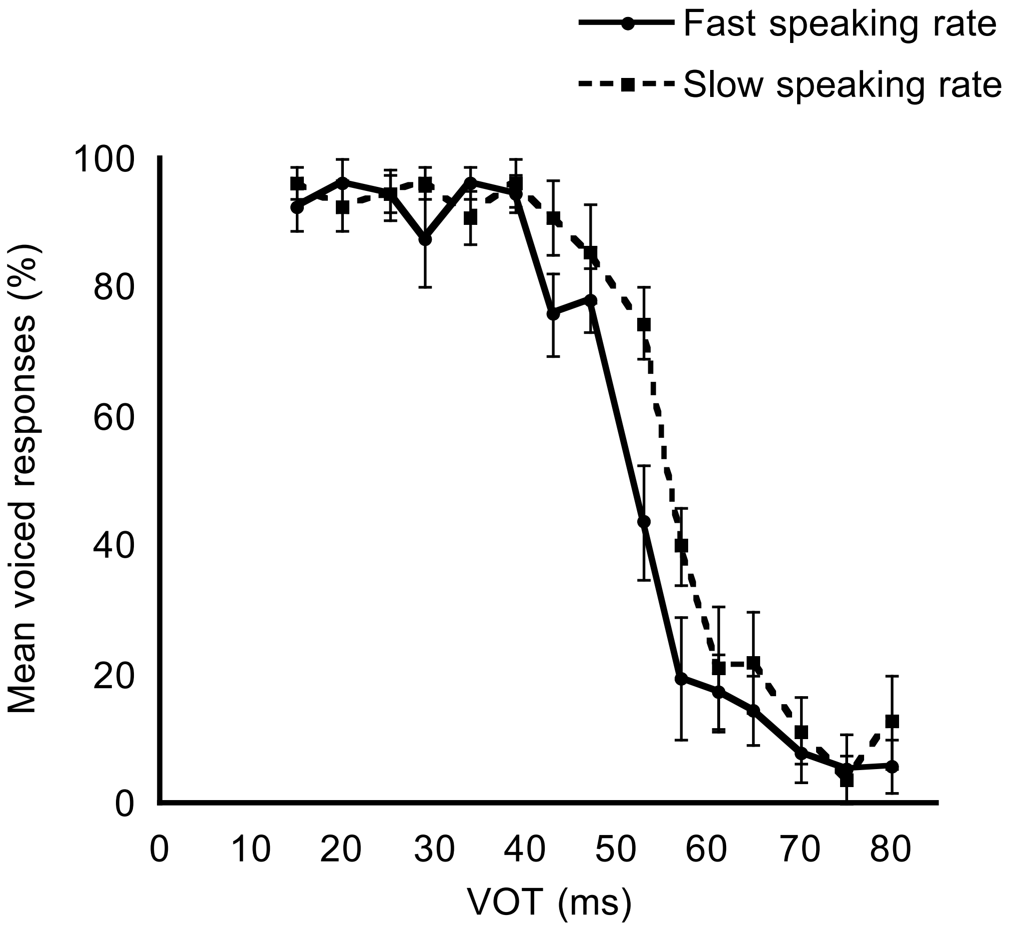

Figure 8. Mean percent voiced responses as a function of voice-onset-time for the fast and slow speaking rates tested in Experiment 2. Error bars indicate standard error of the mean.

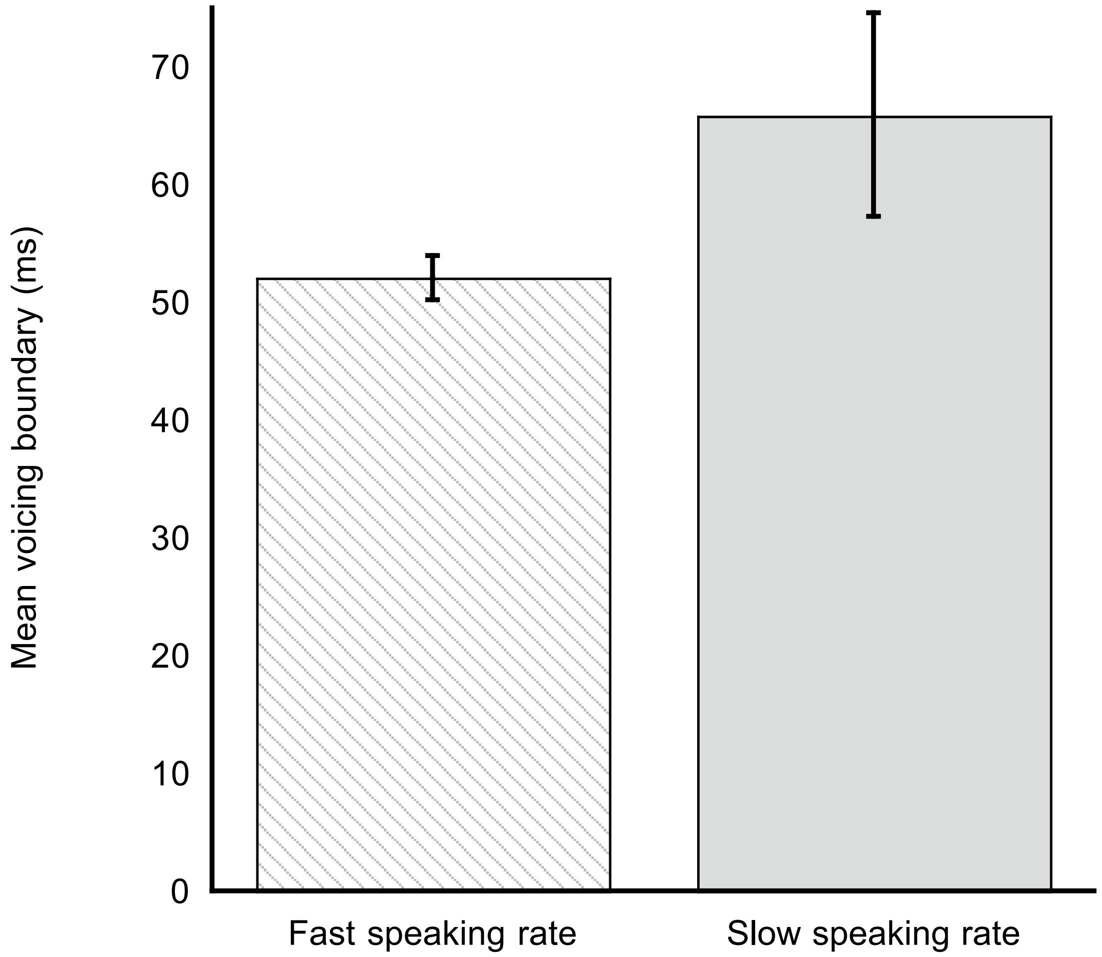

Figure 9. Mean voicing boundary for the children in Experiment 2 for the fast and slow speaking rates tested in Experiment 2. Error bars indicate standard error of the mean.

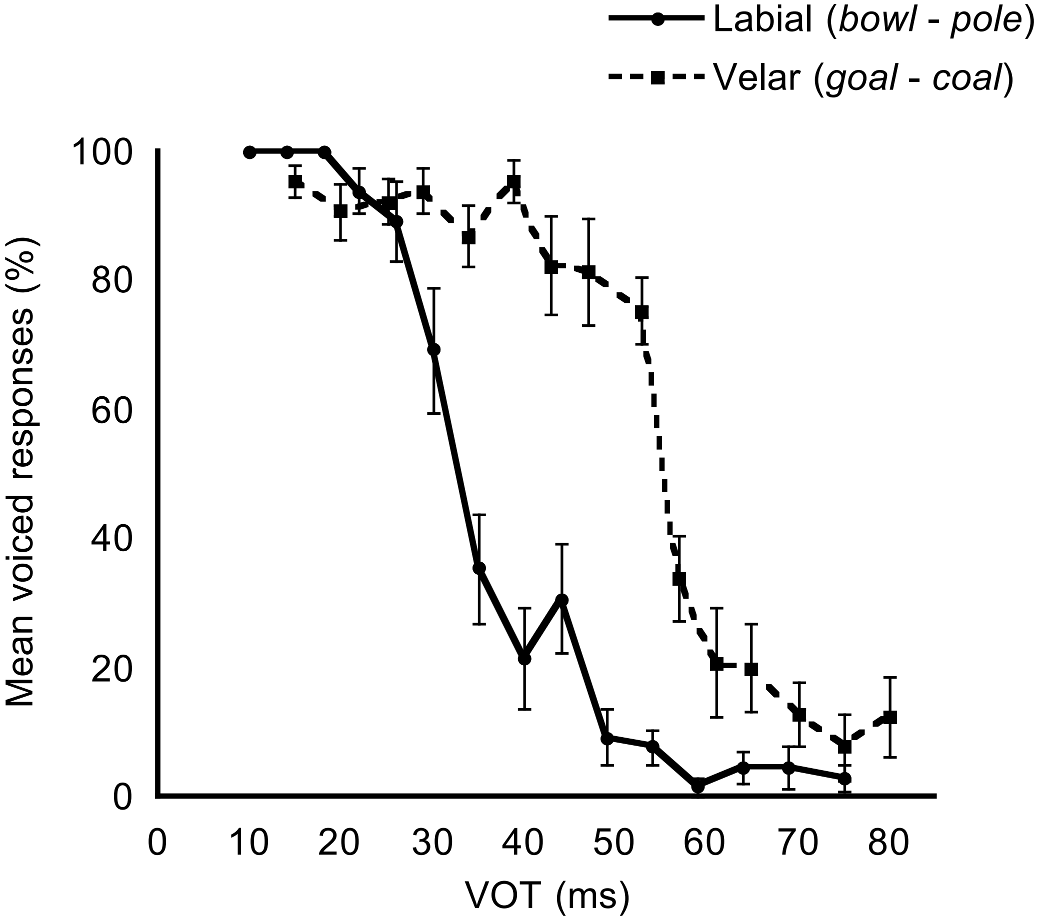

Figure 10. Mean percent voiced responses as a function of voice-onset-time for the labial and velar continua tested in Experiment 2. Error bars indicate standard error of the mean.

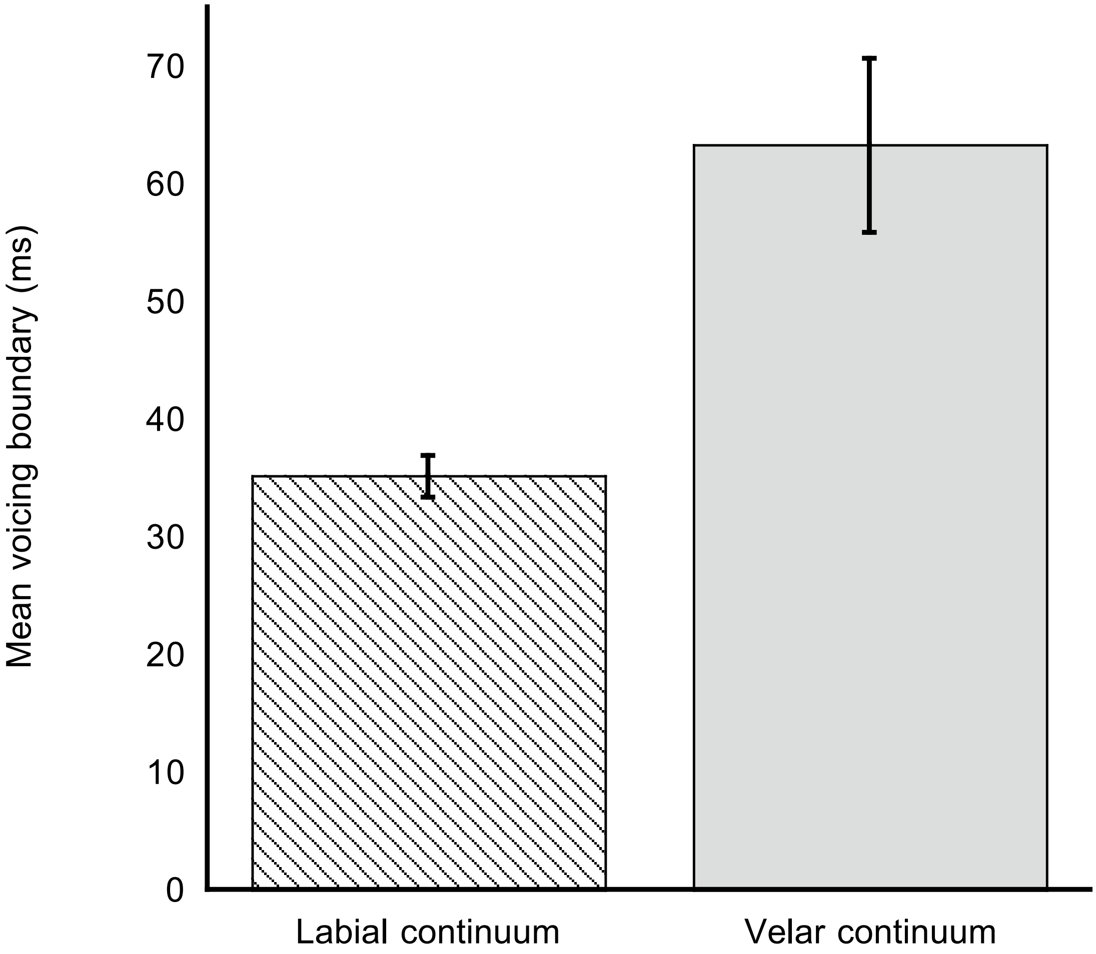

Figure 11. Mean voicing boundary for the children in Experiment 2 for the labial and velar continua tested in Experiment 2. Error bars indicate standard error of the mean.

Tables

Table 1

VOT (ms) and word duration (ms) for the speaking rate continua used in Experiments 1 and 2.

	Fast speaking rate		Slow speaking rate	
Token	VOT	Duration	VOT	Duration
1	15	317	15	478
2	20	317	20	478
3	25	317	25	478
4	29	317	29	478
5	34	317	34	478
6	39	317	39	478
7	43	317	43	478
8	47	317	47	478
9	53	317	53	478
10	57	317	57	478
11	61	317	61	478
12	65	317	65	478
13	70	317	70	478
14	75	317	75	478
15	80	317	80	478

Table 2

VOT (ms) and word duration (ms) for the place of articulation continua used in Experiments 1 and 2.

	Labial continuum		Velar continuum	
Token	VOT	Duration	VOT	Duration
1	10	479	15	478
2	14	479	20	478
3	18	479	25	478
4	22	479	29	478
5	26	479	34	478
6	30	479	39	478
7	35	479	43	478
8	40	479	47	478
9	44	479	53	478
10	49	479	57	478
11	54	479	61	478
12	59	479	65	478
13	64	479	70	478
14	69	479	75	478
15	75	479	80	478

Note. The velar continuum is the same as reported for the slow speaking rate continuum in Table 1.

Appendix A

Pictures used for button labels in the phonetic identification task.

goal

coal

bowl

pole

Appendix B

Form use to collect demographic information for each child.

Subject Information (Child)

1) Your relation to the child ___ Mother ___Father ___Guardian ___Other

2) Child's Gender: ___ Male ___ Female

3) Child's Birth Date (month, day, year): ___________________

4) Was the child's born ______ Full Term or _____ Pre-Term?

5) Place of Birth: ___________________

6) Are you fluent in any language other than English? ___ Yes ___ No

If "Yes", specify: ____

7) Is your child fluent in any language other than English? ___ Yes ___ No

If "Yes", specify: ____

8) What primary language(s) are spoken in the child's home? _______________________

9) Is there a family history of speech, language, or hearing disorders? ___ Yes ___ No

If "Yes", specify: ____

10) Has your child been diagnosed with a speech, language, or hearing disorders?

If "Yes", specify: ____

11) May we contact you for further studies? ___ Yes ___No

Hearing Screening (screen at 20 dB)

RA Initials:

Left ear

Frequency (Hz)	Response (✓/X)
500	
1000	
2000	
4000	

Right ear

Frequency (Hz)	Response (✓/X)
500	
1000	
2000	
4000	

Appendix C

Comprehensive data for children participants sorted by age. A measure of categorical processing (r) is shown for each continuum in addition to age (months) and gender. Standard score (Score) and percentile (%) are provided for measures of articulation, phonology, and non-verbal cognition. Cells with missing data are indicated "n/a."

	r		*r*				Goldman-		Kahn-Lewis		Raven's	
Child	Fast	Slow	Labial	Velar	Age	Gender	Score	%	Score	%	Raw	%
14	**n/a**	**0.31**	**n/a**	**0.31**	60	F	84	11	92	20	n/a	n/a
15	**n/a**	**0.00**	**0.00**	**0.00**	60	F	113	74	n/a	n/a	24	95
20	0.71	0.98	0.98	0.98	60	M	87	21	n/a	n/a	n/a	n/a
12	**0.95**	**n/a**	**0.81**	**n/a**	64	M	59	<1	n/a	n/a	n/a	n/a
2	**0.80**	**n/a**	**n/a**	**n/a**	65	M	n/a	n/a	n/a	n/a	n/a	n/a
9	**n/a**	**0.92**	0.99	0.92	69	F	111	69	108	71	16	50-75
18	0.94	0.72	0.97	0.72	71	M	110	71	108	74	25	95
1	**0.00**	**0.42**	**0.69**	**0.42**	76	M	109	63	107	66	15	50-75
5	0.80	0.96	0.99	0.96	81	M	108	>64	108	>70	26	75-90
16	**0.25**	**0.75**	0.83	0.75	87	M	109	>58	n/a	n/a	34	95
6	0.91	0.77	0.91	0.77	90	F	103	45	104	51	29	90
19	**0.00**	**0.00**	**0.97**	**0.00**	94	M	108	>52	n/a	n/a	30	90
10	0.90	0.97	0.97	0.97	96	M	107	>49	107	>50	28	75-90
7	1.00	0.98	0.99	0.98	98	M	107	>49	107	>50	27	75
8	1.00	0.97	0.97	0.97	98	F	104	>51	106	>49	19	25-50
4	0.98	0.98	0.99	0.98	99	F	104	>48	106	>45	31	90
17	0.98	0.86	0.93	0.86	115	F	102	>39	105	>37	24	25-50
11	0.99	0.99	1.00	0.99	116	M	100	28	104	38	34	90-95
13	**0.70**	**n/a**	**0.81**	**n/a**	116	M	100	24	101	35	30	50-75
3	0.97	0.98	0.94	0.98	128	F	101	>35	104	>32	31	50-75

Note: Bold font is used to indicate participants who were excluded from the results.