The impact of brief restriction to articulation on children's subsequent speech production

Amanda Seidl^{a)} and Françoise Brosseau-Lapré

Speech, Language, & Hearing Sciences, Purdue University, West Lafayette, Indiana 47906, USA

Lisa Goffman

Callier Center for Communication Disorders, Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, Texas 75080, USA

(Received 10 November 2017; revised 2 January 2018; accepted 4 January 2018; published online 9 February 2018)

This project explored whether disruption of articulation during listening impacts subsequent speech production in 4-yr-olds with and without speech sound disorder (SSD). During novel word learning, typically-developing children showed effects of articulatory disruption as revealed by larger differences between two acoustic cues to a sound contrast, but children with SSD were unaffected by articulatory disruption. Findings suggest that, when typically developing 4-yr-olds experience an articulatory disruption during a listening task, the children's subsequent production is affected. Children with SSD show less influence of articulatory experience during perception, which could be the result of impaired or attenuated ties between perception and articulation.

© 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5021710>

[AL]

Pages: 858–863

I. INTRODUCTION

All children experience brief restrictions to their speech articulators on a daily basis in the course of eating, sucking lollipops, or just wrestling with annoying siblings. Yet, we have little understanding of how these brief restrictions impact children's subsequent production of speech sounds. The primary question asked in the present work was whether children rely on articulatory networks during word learning, even during a listening task. Specifically, we asked whether disruption of these networks during perceptual input influences the speech production that follows. While it is obvious that acoustic components of perceptual experience influence production, little is known about the role of covert articulatory experience during learning.

It has been reported that a 6-month-old's own articulatory movements contribute to her perception of sound categories (Bruderer *et al.*, 2015; Majorano *et al.*, 2013), suggesting a system open to articulatory feedback. However, other data suggest that 2-yr-olds, unlike adults, actually show insensitivity to auditory perturbations such that their productions are unaltered in response to formant shifting in their input (MacDonald *et al.*, 2012). While the interaction between perceptual experience and production is a central question in the literature (e.g., Guenther, 2003, 2006, 2016; Guenther and Vladusich, 2012; MacDonald *et al.*, 2012; Ménard *et al.*, 2016), remarkably little work has addressed this issue directly in young children. Further, children with difficulty acquiring sound categories [i.e., diagnosed with speech sound disorders (SSDs)] may be informative regarding links across perception and production.

The purpose of the present work is to evaluate how young children, both typically developing (TD) and with a SSD,

produce speech sounds immediately after experiencing motor restrictions during the perception phase of acquiring a novel word. Children with SSD are particularly relevant to this question because of their documented deficits in both perception (e.g., Rvachew, 2007) and production (e.g., Shriberg and Kwiatkowski, 1994) of speech. Differences in sensitivity in response to perceptual input during word form learning may influence the productions that follow. It is this mapping of perception to production that frames the current study.

We know that most adults can quickly adjust or compensate online for the presence of restrictions by shifting their articulatory movements and hence their formants (e.g., Ménard *et al.*, 2016; Savariaux *et al.*, 1995) and may also show more exaggerated articulation or clearer speech resulting from hyperarticulation after the restriction is removed, at least for certain contrasts (Ménard *et al.*, 2016). For children, there are few studies of the impact of brief restrictions to the articulators; however, data from Ménard *et al.* (2008, 2016) suggest that 4-yr-olds may show a hyperarticulation effect post restriction. Specifically, in Ménard *et al.* (2008), children's production of the vowel [u] was restricted by the presence of a lip tube. While the authors were primarily interested in comparing the children's ability to compensate for the presence of the lip tube with that of adults' by having the children produce [u] with the lip tube in place (see Ménard *et al.*, 2016, for full comparison), they also examined the children's production of [u] after the tube was removed. Results revealed that while, overall, child participants did not appear to compensate for the presence of the lip tube by altering $F1$ and $F2$ while the lip tube was in their mouths, they did show an effect *after* the lip tube was removed. Specifically, $F1$ and $F2$ were both lower (i.e., hyperarticulated) after the tube-restriction was removed.

Given the daily presence of restrictions to children's speech noted above, this hyperarticulation is likely adaptive

^{a)}Electronic mail: aseidl@purdue.edu.

and suggests that, although by 4 years children do not yet show an ability to compensate for disruptions online, children who are typical speech and language learners are not adversely impacted by these daily restrictions since they are still adept at producing the sound contrast *after* the tube is removed and even show hyperarticulation after the restriction (Ménard *et al.*, 2016). This hyperarticulation effect in young learners may be specific to a particular sound and restrictor context or may only occur when the child first produces speech sounds with the restrictor in place (as occurred in Ménard *et al.*, 2016). In addition, it may only hold true for children whose perceptual and articulatory systems are intact (TD) and not for children with SSD who show deficits in speech sound learning. It may be that differences in sensitivity to perceptual input or in linking perception and production in SSD contribute to deficits in acquiring the sound patterns crucial for the accurate production of words.

In this paper, we explore whether a difficult-to-perceive and -produce place of articulation contrast would be enhanced (or disrupted) by the presence of a brief restriction and whether this would hold true for children with SSD as well as those who are TD. Children with SSD are particularly relevant to this question since their speech production deficit may reside in perceptual, phonological, and/or articulatory processing (e.g., International Expert Panel on Multilingual Children's Speech, 2012). Even in the face of perceptual and processing deficits, the clinical hallmark of SSD is at the level of production of sound errors. The objective of this study was first to evaluate whether, as reported by Ménard *et al.* (2008), children who are typical learners are sensitive to articulatory restrictions in productions that immediately follow such restrictions. Second, we asked whether children who are characterized by their speech sound production deficits are differentially sensitive to restrictions in the input. Together, these findings will provide insight into perception-production links during learning.

Methodologically, we chose to examine children's production of /s/ and /f/ for a few key reasons. First, these two sounds are difficult to produce and show protracted development even for children with TD (e.g., Nittrouer *et al.*, 1989; Li *et al.*, 2009), and therefore may be particularly challenging for children with SSD. Second, it is relatively easy to disrupt the typical articulator movements for these two sounds.

It is worthy of note that the children in Ménard *et al.* (2008) may have hyperarticulated the vowel [u] after the lip tube was removed because, despite an inability to compensate online for the presence of the lip tube by altering their formants, perception of their own articulations during the lip tube phase led to this response. In fact, Ménard *et al.* (2008) suggest that the presence of this hyperarticulation may be the first step in compensating for the presence of the lip tube by altering formants. This hypothesis leads to a secondary question we will ask in our study. Specifically, we will explore whether such hyperarticulation also occurs in the absence of feedback from the child's own articulatory system (when the child does not produce sounds with the restrictor in place, but only after it is removed). This allows us to examine whether hyperarticulation occurs solely as the result of the articulators being restricted during the

perceptual component of the learning experience. Findings will reveal whether restrictions of articulatory movement during perceptual experience influence the speech production that follows.

II. METHODS

A. Participants

Forty 4- to 5-yr-olds, all of whom passed a hearing screening and an assessment of nonverbal cognition, participated in the study. Half of these participants were diagnosed with SSD and were an average of 61 months of age [standard deviation (SD) = 5.8; 6 female] and half were TD and were an average of 53 months of age (SD = 4.9; 8 female). These two groups were constructed based on scores on either the Bankson-Bernthal Test of Phonology (BBToP; Bankson and Bernthal, 1990) or the Goldman-Fristoe Test of Articulation-3 (GFTA-3; Goldman and Fristoe, 2015). Specifically, children who received standard scores of less than 85 on either the BBToP or the GFTA-3 were placed in the SSD group. Half of the children in the SSD group were also classified as having a language impairment based on scores on the SPELT-P2 (Dawson *et al.*, 2005) but, given that we did not address a specific hypothesis related to language impairment in this paper, this group was not treated separately from the larger group. Parents were compensated \$10 and participants received a toy for each visit and were provided with consent on their first visit and assent on both visits.

B. Stimuli

The stimuli for the experiment were a series of six colorful novel animal drawings coupled with auditory novel consonant-vowel-consonant (CVC) labels (Fig. 1; Ohala, 1999). All CVCs contained /s/ or /f/ in word-initial position, such that half began with /s/ and half with /f/. All words were controlled for phonotactic probability and neighborhood density (Storkel and Hoover, 2010) and were presented via PowerPoint to children on a large central video monitor in a single-walled sound booth built to resemble a small theatre.

C. Equipment

Children were recorded producing words via a wireless lavalier microphone (AKG WMS40) connected to a digital Marantz PMD 660 recorder. Visual and auditory stimuli were played in PowerPoint over a Macintosh computer and presented on a large video screen in a single-walled sound booth. A silicone fish (item OM8221, Therapy Shoppe; see Fig. 2) advertised for use for pediatric oral motor therapy was connected to a cloth lanyard and used as a restrictor to articulatory movement.

D. Procedure

Children were brought into the sound booth for testing and had a lavalier mic clipped to their clothing. Children were given the option of sitting on a caregiver's lap or alone.

Children were tested on each of 2 days and testing occasions were separated by 1 week. On one day (counterbalanced

FIG. 1. (Color online) Stimuli and their labels.

across participants) children were given the fish. No differences were found between the two groups counterbalanced for day of restrictor. After placing the lanyard around their necks children were instructed on how to place the fish in their mouths. They were told to keep the fish in their mouths except for while they were producing target words. All children complied with this request and no differences were found between the two groups counterbalanced for the day of restrictor. Thus, while children listened to the target words they held the fish in their mouths and only removed it for production of words (Fig. 2). Critically, the fish toy blocked covert or overt tongue movement to the alveolar and palatal places of articulation required for /s/ and /ʃ/. Children heard each word four times during the slide show and thus had the opportunity to produce each word at least four times (most children did so with only a small number of children producing only three repetitions or five repetitions). On the day of testing without the restrictor-fish, all other procedures were the same except for the presence of the restrictor and lanyard.

E. Acoustic analyses

Trained research assistants tagged the onsets and offsets of all productions of /s/ and /ʃ/ from target words at upward zero crossings in Praat (Boersma and Weenink, 2013). Note that sometimes the child produced each word more than once

at a prompt. If this was the case we tagged all productions. Following methods used in Jongman *et al.* (2000), a Praat script was used to extract the highest spectral peak, center of gravity (CoG; the most frequent acoustic measure of place of articulation for fricatives, e.g., Stevens, 1998) using a 40 ms full Hamming window placed in the middle of the friction noise, and skewness. Spectral peak estimation was based on spectra generated by means of fast Fourier transform.

All three measures obtained (Peak, CoG, Skewness) are known to differentiate /s/ and /ʃ/. Specifically, both Peak and CoG should be higher for /s/ than for /ʃ/ since the point of constriction is closer to the lips in /s/ and both, indirectly, provide measures of vocal tract length. If the restrictor led to greater differentiation between the two sounds, it was predicted that differences in Peak and CoG for the two sounds would be exaggerated in the session with the restrictor as compared to the session without the restrictor. Skewness was extracted since it has also been shown to be able to discriminate between sibilant places of articulation (e.g., Jongman *et al.*, 2000; Nissen and Fox, 2005; Nittrouer, 1995) with positive skew occurring with /ʃ/ and negative skew occurring with /s/ related to changes in the length of the cavity. Though none of these measures is perfect in reliably classifying the sibilants under investigation, optimal results can be obtained when CoG and skewness are used together to identify fricative place of articulation (e.g., Jongman *et al.*, 2000; Li *et al.*, 2009).

III. RESULTS

We first ran a linear mixed effects model in JMP statistical software with Peak as the dependent variable and Diagnosis (SSD, TD), Sound (/s, ʃ/), and Condition (Restrictor, No-Restrictor) as independent variables, and Subject as a random effect to control for differences associated with this variable. Here we report results from fixed effects tests. These revealed a main effect of Sound [$\beta_{\text{sound}} = 1326$, standard error (SE) = 65, $p < 0.0001$; /s/ was higher than /ʃ/], an interaction between Diagnosis and Sound ($\beta_{\text{diagnosis} \times \text{sound}} = -496$, SE = 65, $p < 0.0001$), and an interaction of Diagnosis $\times$ Condition $\times$ Sound ($\beta_{\text{diagnosis} \times \text{condition} \times \text{sound}} = 159$, SE = 65, $p < 0.014$). All other main effects and interactions were not significant ($\beta_{\text{diagnosis}} = -275$, SE = 261, $p = 0.298$; $\beta_{\text{condition}} = -102$, SE = 145, $p = 0.438$; $\beta_{\text{diagnosis} \times \text{condition}} = 49$, SE = 145, $p = 0.737$; $\beta_{\text{condition} \times \text{sound}} = -77$, SE = 65, $p = 0.234$). Given the interactions with Diagnosis and our predictions, we ran the same model within each diagnostic group. Results for

FIG. 2. (Color online) Silicone fish which acted as a restrictor for one day of testing and child using restrictor.

FIG. 3. Box plots showing means, range, and SD for Peak $\times$ Condition $\times$ Sound separated by diagnostic group.

the SSD group revealed a main effect of Sound ($\beta_{\text{sound}} = 830$, $\text{SE} = 96$, $p < 0.0001$), but no other main effects or interactions ($\beta_{\text{condition}} = -52$, $\text{SE} = 212$, $p = 0.811$; $\beta_{\text{condition} \times \text{sound}} = 82$, $\text{SE} = 96$, $p = 0.392$). Results for the TD group revealed a main effect of Sound ($\beta_{\text{sound}} = 1821$, $\text{SE} = 86$, $p < 0.0001$), and an interaction of Condition $\times$ Sound ($\beta_{\text{condition} \times \text{sound}} = -236$, $\text{SE} = 86$, $p < 0.0063$), but no main effect of Condition ($\beta_{\text{condition}} = -151$, $\text{SE} = 198$, $p = 0.456$). To explore the source of the interaction, we ran the same model within each condition for children with TD. This revealed that there was a larger effect of Sound with the restrictor ($\beta_{\text{sound}} = 2057$, $\text{SE} = 124$, $p < 0.0001$), than without it ($\beta_{\text{sound}} = 1584$, $\text{SE} = 120$, $p < 0.0001$). Thus, both groups showed higher peaks with /s/ than /ʃ/, but the TD group showed a larger distinction overall and this was driven by a larger difference between the two sounds in the Restrictor than in the No-Restrictor condition. This effect is shown in Fig. 3 with summary data provided in Table I.

We next ran the same mixed model on CoG. Results revealed a main effect of Sound ($\beta_{\text{sound}} = 664$, $\text{SE} = 52$, $p < 0.0001$), a marginal main effect of Diagnosis ($\beta_{\text{diagnosis}} = -509$, $\text{SE} = 256$, $p = 0.0554$), an interaction between Sound and Diagnosis ($\beta_{\text{diagnosis} \times \text{sound}} = -299$, $\text{SE} = 52$, $p < 0.0001$), and an interaction of Diagnosis and Condition ($\beta_{\text{diagnosis} \times \text{condition}} = -133$, $\text{SE} = 52$, $p < 0.0103$), but no other main effects or interactions ($\beta_{\text{condition}} = -5$,

$\text{SE} = 52$, $p = 0.93$; $\beta_{\text{condition} \times \text{sound}} = -52$, $\text{SE} = 52$, $p = 0.31$; $\beta_{\text{diagnosis} \times \text{condition} \times \text{sound}} = 65$, $\text{SE} = 52$, $p = 0.21$). Given the marginal main effect of Diagnosis and the interactions with Diagnosis, we ran the same mixed model within each diagnostic group. Results for the SSD group revealed a main effect of Sound ($\beta_{\text{sound}} = 365$, $\text{SE} = 75$, $p < 0.0001$), but no other main effects or interactions ($\beta_{\text{condition}} = -137$, $\text{SE} = 75$, $p = 0.068$; $\beta_{\text{condition} \times \text{sound}} = 13$, $\text{SE} = 75$, $p = 0.861$). Results for the TD group revealed a main effect of Sound ($\beta_{\text{sound}} = 964$, $\text{SE} = 71$, $p < 0.0001$), and no other main effects or interactions ($\beta_{\text{condition}} = -128$, $\text{SE} = 71$, $p = 0.072$; $\beta_{\text{condition} \times \text{sound}} = -117$, $\text{SE} = 71$, $p = 0.099$). Thus, both groups showed higher CoG for /s/ than /ʃ/, but the effect was larger in the TD group, as shown in Fig. 4.

Finally, we ran the same mixed model on Skewness. Results revealed main effects of Sound ($\beta_{\text{sound}} = -0.5$, $\text{SE} = 0.036$, $p < 0.0001$) and Diagnosis ($\beta_{\text{diagnosis}} = 0.5$, $\text{SE} = 0.23$, $p < 0.02$). There was also an interaction between Diagnosis and Sound ($\beta_{\text{diagnosis} \times \text{sound}} = 0.2$, $\text{SE} = 0.036$, $p < 0.0001$) and Condition and Sound ($\beta_{\text{condition} \times \text{sound}} = 0.07$, $\text{SE} = 0.036$, $p < 0.03$). All other main effects and interactions were not significant. Given the main effect of Diagnosis and the interaction of Diagnosis with Sound, we ran the same model within each diagnostic group. Results for the SSD group revealed no main effects or interactions ($\beta_{\text{sound}} = -0.2$, $\text{SE} = 0.10$, $p = 0.055$; $\beta_{\text{condition}} = 0.5$, $\text{SE} = 0.23$, $p = 0.624$; $\beta_{\text{condition} \times \text{sound}} = 0.02$, $\text{SE} = 0.053$, $p = 0.7$). Results for the TD group revealed a main effect of Sound ($\beta_{\text{sound}} = -0.66$, $\text{SE} = 0.11$, $p < 0.0001$) and an interaction of Condition and Sound ($\beta_{\text{condition} \times \text{sound}} = 0.14$, $\text{SE} = 0.04$, $p < 0.0014$), but no effect of Condition ($\beta_{\text{condition}} = -0.08$, $\text{SE} = 0.14$, $p = 0.538$). The same model within condition for TD revealed a larger effect of Sound with the restrictor ($\beta_{\text{sound}} = -0.80$, $\text{SE} = 0.12$, $p < 0.0001$) than without the restrictor ($\beta_{\text{sound}} = -0.5$, $\text{SE} = 0.12$, $p < 0.0007$), as shown in Fig. 4. This suggests that Skewness was distinct for the two sounds for TD children regardless of restrictor, but that sounds may have been hyperarticulated post-restrictor. However, children with SSD did not appear to use skewness (i.e., manipulation of length of the cavity) as a mechanism to differentiate the two sounds.

IV. DISCUSSION

Regardless of restrictor, both diagnostic groups produced voiceless sibilant sounds which were distinct. While the effect of Sound was, over all measures, smaller in children with

TABLE I. Summary data for all acoustic measures with means and SDs separated by sound, condition, and diagnosis.

Restrictor	Diagnosis	Sound	Mean Peak	SD Peak	Mean Skew	SD Skew	Mean CoG	SD CoG
No-Restrictor	SSD	s	7600.69	3727.43	1.53	2.39	4683.15	2926.65
No-Restrictor	SSD	sh	5772.33	3042.16	1.95	2.28	3950.18	2513.79
No-Restrictor	TD	s	8520.46	3081.38	0.26	1.43	6294.73	2705.25
No-Restrictor	TD	sh	5335.93	2604.88	1.33	1.64	4588.07	1855.50
Restrictor	SSD	s	7383.32	3634.47	1.24	2.23	5038.99	2888.62
Restrictor	SSD	sh	5882.64	3422.21	1.70	2.31	4354.73	2551.75
Restrictor	TD	s	9247.51	2967.76	0.16	1.76	6284.38	2779.86
Restrictor	TD	sh	5146.50	2821.76	1.79	1.82	4110.59	1928.53

FIG. 4. (Color online) Box plots showing means, range, and SD for $\text{CoG} \times \text{Condition} \times \text{Sound}$ and $\text{Skewness} \times \text{Condition} \times \text{Sound}$ separated by diagnostic group.

SSD, they nonetheless as a group showed an ability to produce distinct /s/ and /ʃ/ categories based on CoG and Peak. The largest difference between the two diagnostic groups was in comparing the restricted with the not restricted condition for Peak and Skewness: The SSD group was unaffected by learning the words with the restrictor present, while for the TD group, the restrictor produced an effect similar to that seen in Ménard and colleagues (2008) when the restrictor was removed for each subsequent production. In sum, our findings reveal that by 4 years of age, children who are typical speech and language learners are highly sensitive to cross-modal input during novel word form learning and modify production even based on brief restriction of articulatory movement during perception. However, children with SSD did not show the same sensitivity to an articulatory restriction during perceptual input and did not modify cues in production following the disruption of potential movement during perception.

Why were children with SSD unaffected by their experience with the restrictor during the listening phase of the task and why were those with TD affected by the restrictor for Peak and Skewness? There are two possible explanations. First, it may be that the children with TD experienced a sustained biomechanical response in response to the perturbed placement of the tongue in relation to the alveolar ridge and palate during the restriction that occurred during the perceptual listening phase. If this were the case, our findings may be related to lower level articulatory phenomena. The distribution of the specific cues affected is not consistent with this explanation. Given that we see this effect only in Peak and Skewness, it seems unlikely that the phenomenon we see here is related to such low-level biomechanical factors. In fact, CoG is an acoustic correlate that is related to place of articulation (Stevens, 1998). Since place was overtly perturbed after the restrictor, but CoG was not, maintenance of a biomechanical response to the restrictor seems an unlikely explanation for our findings. Further, there is no reason to

expect that biomechanical factors would have differential effects across diagnostic groups, since our participants with SSD completed an oral-speech mechanism examination and do not show low-level oral motor deficits, nor is SSD severity typically associated with motor skill (Lewis *et al.*, 2011). Therefore, this explanation would not be supported by the results which show differential effects of the restrictor for the two diagnostic groups.

Another possible explanation is that, through somatosensory information, articulatory restriction influences higher order perception-production networks. Data from adults suggest that altering somatosensory information during a perceptual task impacts speech perception. For example, Ito *et al.* (2009) found that stretching the facial skin during a listening task impacted how sounds were perceived. Similarly, in our task children with TD experienced sound contrasts that were altered due to somatosensory feedback during listening in the restrictor condition; this resulted in a greater differentiation between the acoustic cues in the two sounds during subsequent production. On this account, the lack of impact of restrictor on children with SSD may be related to an impairment in somatosensory feedback during perception. Thus, this task and these results may point to differences in children with SSD's abilities to link perceptual representations with articulatory actions. One supporting piece of evidence for this account may be provided by other works that suggest perceptual differences in children with SSD (e.g., Edwards *et al.*, 2002; Rvachew and Grawburg, 2006). Thus, both pure perception deficits and deficits in production-perception links may contribute to explain the lack of greater differentiation of sounds post-restrictor in children with SSD.

Children with SSD are defined by their deficits in producing stable and accurate articulatory configurations. According to many developmental models of speech perception-production, the achievement of the targeted vocal

tract shape requires the integration of perceptual and somatosensory streams (Bruderer *et al.*, 2015; Guenther, 2016; Ménard *et al.*, 2016). For example, according to the DIVA model, speakers plan movement trajectories in acoustic space—feed-forward commands which contain information regarding the position and the velocity of the articulators, allowing the speaker to achieve the desired vocal tract shape (Guenther, 2016). A feedback control system is integrated in the model such that speakers can build corrective feedback commands when the production of the target does not match acoustic or sensory expectations. It seems likely that the lack of greater differentiation of sounds in children with SSD indicates deficits in the control mechanisms integrating sensorimotor (or vocal tract) and perceptual components of speech production. Bruderer and colleagues (2015) suggest that sensorimotor deficits may have downstream effects on perception and language, another indicator of the critical linkages that may disrupt multiple dimensions of speech and language learning in children with SSD.

Together, these findings provide insights into the perception-production links underlying the acquisition of novel word forms. Longstanding theories have considered influences not only of perception on production, but also the reverse. In the 1960s, it was proposed via the motor theory of speech perception (Liberman *et al.*, 1967) that perceiving speech is perceiving articulatory gestures. The present work, consistent with that of Ménard *et al.* (2008), provides evidence that the motor system alters how we perceive and subsequently produce speech ourselves. Thus, the data from this paper provide a key piece of evidence showing that motor restriction during speech perception can immediately impact subsequent production and further unifies speech perception and sensorimotor experience with articulation.

ACKNOWLEDGMENTS

A.S., L.G., and F.B. contributed to the initial planning of the study. A.S. and L.G. designed the experiment. A.S., L.G., and F.B. supervised data collection. A.S. and L.G. created approaches to coding. A.S. supervised acoustic analysis and conducted statistical analyses and prepared the graphics. A.S., L.G., and F.B. co-wrote the manuscript. The authors thank Katie Isbell, Lakin Brown, and Rana Abu Zhaya for their assistance in collecting and analyzing the data. Portions of this work were supported by the National Institute on Deafness and Other Communication Disorders Grant No. R01 DC04826. This work was performed after obtaining IRB approval.

- Bankson, N. W., and Bernthal, J. E. (1990). *Bankson-Bernthal Test of Phonology* (Riverside Press, Chicago, IL).
- Boersma, P., and Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.5) [Computer program], online (Last viewed May 12, 2017).
- Bruderer, A., Danielson, D., Kandhadai, P., and Werker, J. (2015). "Sensorimotor influences on speech perception in infancy," *Proc. Natl. Acad. Sci. U.S.A.* **112**(44), 13531–13536.
- Dawson, J., Eyer, J. A., and Fonkalsrud, J. (2005). *Structured Photographic Expressive Language Test—Preschool*, 2nd ed. (Janelle Publications, DeKalb, IL).

- Edwards, J., Fox, R., and Rogers, C. (2002). "Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy," *J. Speech, Lang., Hear. Res.* **45**, 231–242.
- Goldman, R., and Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation*, 3rd ed. (Pearson, Bloomington, MN).
- Guenther, F. H. (2003). "Neural control of speech movements," in *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, edited by A. Meyer and N. Schiller (Mouton de Gruyter, Berlin, Germany), pp. 209–239.
- Guenther, F. H. (2006). "Cortical interactions underlying the production of speech sounds," *J. Commun. Disorders* **39**, 350–365.
- Guenther, F. H. (2016). *Neural Control of Speech* (MIT Press, Cambridge, MA).
- Guenther, F. H., and Vladusich, T. (2012). "A neural theory of speech acquisition and production," *J. Neuroling.* **25**(5), 408–422.
- International Expert Panel on Multilingual Children's Speech. (2012). "Multilingual children with speech sound disorders: Position paper," <http://www.csu.edu.au/research/multilingual-speech/position-paper> (Last viewed May 12, 2017).
- Ito, T., Tiede, M., and Ostry, D. (2009). "Somatosensory function in speech perception," *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1245–1248.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**(3), 1252–1263.
- Lewis, B. A., Avrich, A. A., Freebairn, L. A., Taylor, H. G., Iyengar, S. K., and Stein, C. M. (2011). "Subtyping children with speech sound disorders by endophenotypes," *Topics Lang. Disorders* **31**(2), 112–127.
- Li, F., Edwards, J., and Beckman, M. E. (2009). "Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers," *J. Phonetics* **37**(1), 111–124.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). "Perception of the speech code," *Psych. Rev.* **74**, 431–461.
- MacDonald, E., Johnson, E., Forsythe, J., Plante, P., and Munhall, K. (2012). "Children's development of self-regulation in speech production," *Current Biol.* **22**, 113–117.
- Majorano, M., Vihman, M., and Depaolis, R. (2013). "The relationship between infants' production experience and their processing of speech," *Lang. Learn. Develop.* **10**, 179–204.
- Ménard, L., Perrier, P., and Aubin, J. (2016). "Compensation for a lip-tube perturbation in 4-year-olds: Articulatory, acoustic, and perceptual data analyzed in comparison with adults," *J. Acoust. Soc. Am.* **139**(5), 2514–2531.
- Ménard, L., Perrier, P., Aubin, J., Savariaux, C., and Thibeault, M. (2008). "Compensation strategies for a lip-tube perturbation of French [u]: An acoustic and perceptual study of 4-year-old children," *J. Acoust. Soc. Am.* **124**(2), 1192–1206.
- Nissen, S. L., and Fox, R. A. (2005). "Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective," *J. Acoust. Soc. Am.* **118**, 2570–2578.
- Nittrouer, S. (1995). "Children learn separate aspects of speech production at different rates: Evidence from spectral moments," *J. Acoust. Soc. Am.* **97**, 520–530.
- Nittrouer, S., Studdert-Kennedy, M., and McGowan, R. S. (1989). "The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults," *J. Speech, Lang., Hear. Res.* **32**, 120–132.
- Ohala, D. K. (1999). "The influence of sonority on children's cluster reductions," *J. Commun. Disorders* **32**(6), 397–422.
- Rvachew, S. (2007). "Phonological processing and reading in children with speech sound disorders," *Am. J. Speech-Lang. Pathology* **16**, 260–270.
- Rvachew, S., and Grawburg, M. (2006). "Correlates of phonological awareness in preschoolers with speech sound disorders," *J. Speech, Lang., Hear. Res.* **49**, 74–87.
- Savariaux, C., Perrier, P., and Orliaguet, J. P. (1995). "Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production," *J. Acoust. Soc. Am.* **98**(5), 2428–2442.
- Shriberg, L. D., and Kwiatkowski, J. (1994). "Developmental phonological disorders I: A clinical profile," *J. Speech Hear. Res.* **37**, 1100–1126.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
- Storkel, H. L., and Hoover, J. R. (2010). "An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English," *Behav. Res. Methods* **42**, 497–506.