The effect of somatosensory input on word recognition in typical children and those with Speech Sound Disorder

Rana Abu-Zhaya[1], Lisa Goffman[2], Françoise Brosseau-Lapré[3], Elizabeth Roepke[4], Amanda Seidl[3]

[1] School of Psychology, University of Plymouth

[2] Callier Center for Communication Disorders, Behavioral and Brain Sciences, University of Texas at Dallas

[3] Department of Speech, Language, and Hearing Sciences, Purdue University

[4] Speech, Language, and Hearing Sciences, Saint Louis University

Author Note

We have no known conflict of interest to disclose.

Correspondence concerning this manuscript should be addressed to Rana Abu-Zhaya, ranaabuzhaya@gmail.com.

Abstract

Purpose: Recent work suggests that speech perception is influenced by the somatosensory system and that oral sensorimotor disruption has specific effects on the perception of speech both in infants who have not yet begun to talk and in older children and adults with ample speech production experience; however, we do not know how such disruptions affect children with speech sound disorders (SSD). Response to disruption of would-be articulators during speech perception could reveal how sensorimotor linkages work for both typical and atypical speech and language development. Such linkages are crucial to advancing our knowledge on how both typically developing and atypically developing children produce and perceive speech.

Method: Using a looking-while-listening task, we explored the impact of a sensorimotor restrictor on the recognition of words whose onsets involve late developing sounds (s, ʃ) for both children with typical development (TD) and their peers with SSD.

Results: Children with SSD showed a decrement in performance when they held a restrictor in their mouths during the task, but this was not the case for children with TD. This effect on performance was only observed for the specific speech sounds blocked by the would-be articulators.

Conclusion: We argue that these findings provide evidence for altered perceptual-motor pathways in children with SSD.

Introduction

Restrictions to articulators while listening to speech are commonly experienced by infants (e.g., sucking a pacifier while their mother speaks to them) and adults (e.g., eating a sandwich during a Zoom meeting). Such restrictions, which modulate somatosensory input, have been shown to impact speech perception under experimental conditions both in adults (e.g., Galantucci et al., 2006; Gick & Derrick, 2009; Ito et al., 2009) and infants (e.g., Bruderer et al., 2015). While some of the results from these studies show only weak effects, findings nonetheless suggest that both motor and auditory systems may be activated during speech perception (e.g., Masapollo & Guenther, 2019; Wilson et al., 2004; Yeung & Werker, 2013), and that speech perception is a broad multimodal and multi-system phenomenon (Gick & Derrick, 2009). These ideas have been addressed in speech perception research for many decades, starting with the motor theory of speech perception –roughly the idea that there is a predictable relationship between acoustic signals and articulatory gestures (Liberman et al., 1967; Liberman & Mattingly, 1985)– and building into more recent theoretical models involving feedback such as The Directions into Velocities of Articulators (DIVA) neural framework (Guenther, 2006; Guenther & Vladusich, 2012; Tourville & Guenther, 2011). Although these theories contribute to explaining the links between speech perception and production, we still do not have a clear understanding of the limits of such links. Specifically, we do not yet know how somatosensory disruptions impact speech processing in the wild, and how resilient different listeners are (e.g., children with speech sound disorder) in processing speech when their articulators are disrupted.

The involvement of the somatosensory system in speech perception may be observed in infants well before they attain fully formed phonological systems and ample experience with articulation: articulatory restrictions have been reported to influence speech perception both at the onset of babbling (at 6 months; Bruderer et al., 2015) and even in the pre-babbling period (at 3 months; Choi et al., 2021). For example, Bruderer and colleagues (2015) reported that a temporary restriction of tongue tip movement interfered with 6-month-olds’ sensitivity in perceiving non-native contrasts that are differentiated by the location of the tongue tip (note that such early perceptual sensitivity to non-native contrasts is typically observed in the first half year of life, e.g., Werker & Tees, 1984). These results were replicated with another group of 6-month-olds, and extended to the perception of a native contrast (/ba/ vs. /da/; Choi et al., 2019). Following up on these findings, Choi et al. (2021) tested pre-babbling 3-month-olds and examined whether neural responses to phonetic contrasts are impacted by oral-motor restrictions to tongue tip movements that are essential to producing these phonemes. Their results showed that restricting tongue tip movements in infants who are not yet capable of producing these sounds was associated with a specific component of phonetic processing rather than broadly disrupting speech perception. Similarly, studies conducted with adults showed that specific modulations to somatosensory input led to specific, rather than broad influences on speech perception, affecting only the speech sounds that are impacted by the somatosensory modulations or would-be articulators (e.g., Gick & Derrick, 2009; Ito et al., 2009; Masapollo & Guenther, 2019). Taken together, this body of literature suggests that sensorimotor experiences may affect speech perception, and that bidirectional auditory-to-motor mappings between articulation and perception may be in place even before infants gain much experience with producing speech sounds. These results provide a framework for studying children with speech sound disorder (SSD), and may have unique implications for these children.

Children with SSD show deficits in speech production, and often, corresponding deficits in speech perception (e.g., they have difficulty perceiving distinctions they cannot produce;

Rvachew & Jamieson, 1989; McAllister Byun, 2012; Sénéchal et al., 2004). Thus, these children serve as an important test case of auditory-motor mapping and the somatosensory influences on speech perception. Further, examining such links in children with SSD may shed light on the nature of the disorder. SSD is diagnosed when a child produces more speech sound errors than same-aged peers in the absence of neurological, psychological, or sensory deficits (Gierut, 1998). SSDs manifest as speech sound production errors in young children, including substitutions or omissions of target speech sounds or syllables. These errors may affect classes or patterns of sounds in production (e.g., all fricatives become stops; Gierut, 1998). Children with SSD also show difficulties in speech perception (e.g., Hearnshaw et al., 2018, 2019; International Expert Panel on Multilingual Children's Speech, 2012), even for speech sounds they produce accurately (e.g., Brosseau-Lapré et al., 2020; Edwards et al., 2002). Further, unlike toddlers with typical speech and language development who are sensitive to mispronunciations of known words (e.g., "baby" produced as "vaby"; Swingley & Aslin, 2000), preschoolers and young school-age children with SSD present with significantly poorer sensitivity to mispronunciations relative to children with TD (Hearnshaw et al., 2019) even in tasks where mispronunciation results in lexical confusion (e.g., "save" vs. "shave", Roepke & Brosseau-Lapré, 2019). Given that SSD is characterized by deficits in production, with frequent co-occurrence of deficits in perception (both broad and specific to the sounds they mispronounce), it is likely that the mapping between articulatory, phonological, and auditory-perceptual components would also be disrupted.

The mapping between articulatory features and perceptual phonological representations can be examined through disrupting somatosensory input. Given that disruptions to the movement of articulators during speech perception (e.g., drinking water while listening to a lecture, sharing a meal with a table of talkative friends) accompany us from infancy, mostly without noticeably impairing our speech perception, it is likely that we develop the necessary skills to compensate for these disruptions. Such compensations are possible due to feedforward and feedback control sub systems that operate during speech perception and production (Guenther, 2006; Guenther & Vladusich, 2012; Tourville & Guenther, 2011). The DIVA neural framework proposes an explanation for the interactions among brain regions that are activated during speech production and perception. Simply put, this framework emphasizes the importance of integrating auditory, motor, and somatosensory information to produce speech accurately (Guenther, 2006; Guenther & Vladusich, 2012; Tourville & Guenther, 2011). Guenther and colleagues propose that the speaker develops an internal model of the correspondence between the shape of the vocal tract and acoustic signals via mapping multimodal information from the movements of articulators (e.g., tactile, proprioceptive, auditory) to specific neural representations. When planning movement trajectories to achieve the target vocal tract shape during speech production, feedforward commands are specified with regards to the position and the velocity of the articulators. A feedback control system allows the speaker to generate corrective articulatory commands when the produced target does not correspond to the expected acoustic or sensory product (Guenther, 2016). Specifically, as a speech sound is produced, the system learns a somatosensory target (i.e., the multimodal sensations) for that sound, which is used during somatosensory feedback control to detect errors related to somatosensory input (e.g., disruption to movement of articulators); following on that process, feedforward commands are corrected and fine-tuned after every production.

The DIVA model emphasizes learning mappings in the developmental context of canonical babbling, suggesting that, “the model first learns the relationship between motor commands and their sensory consequences during a process analogous to infant babbling” (Tourville & Guenther, 2011, p. 9). Atypical child speech learning has also been modeled, with Terband and colleagues (2014a) showing that auditory processing deficits that occur in conjunction with motor processing deficits result in speech sound distortions and token-to-token variability (i.e., features of speech production consistent with childhood apraxia of speech). These modeling studies suggest that auditory self-monitoring is important for some speech production disorders. Interestingly, phonological deficits were deemed to have a motor rather than an auditory processing origin. While these modeling findings are provocative, there is little work assessing the relationship between sensory monitoring and aspects of production or perception in young children with TD or SSD (see for example Terband et al., 2014b).

Our aim here is to determine whether and how young children with TD and their peers with SSD accommodate disrupted somatosensory input during speech perception. This work relies on the assumption that there is an interaction between perception and production systems, and is supported by theoretical models like DIVA (Guenther, 2016) and the motor theory of speech perception (Liberman et al., 1967). Previous studies show that *some* children with SSD have more difficulties than children with TD in compensating for the effect of a bite-block when *producing* vowels in CVC words (e.g., *heat, hit,* and *hat*; Edwards, 1992) and in bisyllabic words (Towne, 1994). Using a different paradigm, one motivated by the infant work of Bruderer and colleagues (2015), Seidl et al. (2018) showed that restricting tongue movements during perception did *not* impact the subsequent production of speech sounds for children with SSD but did for children with TD. Specifically, in two experimental conditions, either while holding an articulatory restrictor in their mouths or not, children heard novel words like "**s**ab" or "**sh**ope" that were mapped to images of novel creatures. After this they were asked to remove the restrictor and repeat the novel words (this was essentially a novel word repetition task; note that this study *did not* directly test whether children mapped those novel words to the novel creatures). Seidl et al. (2018) examined the distinctness of the acoustics of the two voiceless fricative onsets produced by children both when listening occurred with and without the restrictor and found that, while children with SSD did not show any difference in production of distinctness of the two sounds when they heard these novel words with a restrictor, children with TD did (they showed a greater difference between the two sounds immediately after removing the restrictor, i.e., hyper articulation). Given the difference between groups (SSD, TD) and the cues implicated (peak, skewness, but not center of gravity), it is unlikely that this difference was due to biomechanical differences between restrictor conditions (with or without), but was likely related to differences in higher order perception-production networks. In short, these results may suggest a deficit in somatosensory feedback in children with SSD or the mapping between somatosensory feedback and representations.

This previous work focused on the production of novel words, which are known to increase processing load. In the current work, we examine speech perception *alone* using familiar words and a paradigm that involves low cognitive load (looking-while-listening) to evaluate whether somatosensory disruption of the articulators affects speech perception (specifically word recognition) in children with SSD and TD. If speech production disorders such as SSD lead to underdeveloped internal models of the correspondence between the shape of the vocal tract and acoustic signals, and underdeveloped feedback and feedforward systems, then, in line with the DIVA model, we predict that providing disruptive somatosensory information during a speech perception task could lead to inaccurate tuning of details of the acoustic signal and feedforward commands that inform future productions of speech sounds. Such inaccuracies may interfere with the ability of children with SSD to compensate for disruptions to would-be articulators in speech perception, thereby negatively influencing their speech perception skills.

But is the impact of such disruptions to somatosensory input specific or broad? Would disruptions to the movements of articulators affect only the specific speech sounds directly influenced by such restrictions or would the impact be broader? Evidence from TD infants seems to suggest that disruption to specific articulators influenced perception of segments linked to the region of would-be articulation but did not impact perception of speech sounds linked to different regions[1] (Bruderer et al., 2015; Choi et al., 2021; similar findings have been reported in adult studies; Gick & Derrick; Ito et al., 2009; Masapollo & Guenther, 2019). For example, the onset of “apple” and “cup” would not be influenced by an alveolar-coronal restrictor, whereas the onset of “sheep” and “sun” would be. In the case of children with SSD, it is unclear whether disruption of somatosensory input would lead to a broad or specific effect on speech perception; however, the literature reviewed above can help us speculate on the nature of such an effect. Specifically, based on developmental work, altered somatosensory input may disrupt speech perception in TD children impacting only speech sounds whose production is influenced by the disruption, but not speech sounds that are not influenced by the disruption. As for children with SSD, it is possible that they would experience a similarly *specific* impact on speech perception, but that impact may be more pronounced than that observed in TD children, thereby leading to a larger disadvantage in speech perception. Alternatively, children with SSD may experience the impact of somatosensory disruptions in a non-specific/broad manner. Under this broad account, the specific speech sounds perceived, or would-be articulators restricted do not matter: Children with SSD would demonstrate poorer perception of all speech sounds when somatosensory input is disrupted, regardless of what sounds are implicated by the disruption.

---

[1] Though note that these previous studies tested the specificity of the restrictors and somatosensory disruptions in *different* tasks which were analyzed using different statistical models; this could limit the strength of conclusions we make about the specificity of these effects.

In this study, we examine how children with TD and SSD perceive words with correctly pronounced and mispronounced coronal sibilants (s/ʃ), both with and without a coronal articulatory restrictor. Before outlining our hypotheses, we note that the premise of this investigation is based on a small, but growing body of literature conducted mostly with infants and adults. Yet, it is still unclear whether these previous results fall on the same spectrum or should be interpreted in the same vein. These previously reported effects may represent limited phenomena in infants and adults under specific experimental conditions that would not necessarily generalize to children or to speech processing in the wild. Hence, it is possible that children's speech perception skills are not impacted by somatosensory disruptions (e.g., the presence of an object in the mouth). That said, we think that the frequency with which children perform speech perception under conditions of somatosensory disruption, and the implications of such disruptions on the speech perception skills of children with SSD warrant the investigation we undertake here.

Our first set of hypotheses relates to speech perception abilities in the absence of a coronal restrictor. Under normal listening conditions, we expect that children with TD will look to the target word longer when the word is correctly pronounced than when it is mispronounced. We expect that children with SSD will have poorer perception of the s-ʃcontrast, showing similar looking times to the target for both the correct pronunciation and the mispronunciation. These findings would corroborate previous findings that children with SSD have poorer speech perception than children with TD, especially for relatively challenging and late developing speech sounds that they are likely to have had difficulty acquiring (Hearnshaw et al., 2019).

Our second set of predictions relates to the effect (or lack thereof) of the coronal restrictor on speech perception. We predict that performance of children with TD would corroborate previous findings with infants and show specific sensitivity to sensorimotor disruption, leading to poorer differentiation between correct and mispronounced coronals with the coronal disruptor in place, and no effect of the restrictor on the perception of non-coronal sounds. For children with SSD, we make different predictions based on potential deficits affecting somatosensory processing or mapping across auditory and production levels during speech perception for these children. If children with SSD are insensitive to sensorimotor cues compared to TD children, we might expect no effect of a restrictor in any phonetic context. However, if children with SSD are globally sensitive to a sensorimotor disruption, then we would expect poorer performance on all targets, whether coronal or foils, with an articulatory disruptor than without. Another possibility is that children with SSD will demonstrate a specific impairment in mapping sensorimotor cues to speech sounds (as suggested in Seidl et al., 2018). In this case, we would expect poorer performance for both the correct and the mispronounced coronal targets in the presence of the coronal restrictor, but no disruption in performance for foils with non-coronal onsets.

Method

Participants

Forty children ages 48 to 74 months participated in the study, which was approved by Purdue University’s Institutional Review Board (1407015009). Participants were recruited through the M.D. Steer Speech and Language Clinic and through local preschools. Caregivers provided informed written consent and children provided verbal assent prior to participating in the study. Our sample included 20 children with SSD with an average age of 61 months in both visits (range = 53-74 months; *SD* = 5.6; 6 female), and an equal number of children with TD who were significantly younger (Wilcoxon’s test: *S* = 974.5; *Z* = -6.234; *p* < .0001) and had an average age of 52 months (range = 48-70 months; *SD* = 4.99; 8 female). All participants passed a hearing screening at 500, 1000, 2000, and 4000 Hz at 20dB bilaterally and had no history of medical conditions such as neurological impairment, autism spectrum disorder, or global developmental delay.

Children were assigned to the SSD and TD groups based on their case history and clinical assessment by a speech-language pathologist. All children with SSD were receiving speech-language therapy or their parents had expressed concerns regarding their child's communication development and their child had been referred for an assessment. All children with SSD also obtained a standard score below 85 on a norm-referenced single word test of articulation (*M* = 64.00, *SD* = 13.72, range = 40-82). For 14 of the children with SSD this measure was the Goldman-Fristoe Test of Articulation, Third Edition (GFTA-3; Goldman & Fristoe, 2015) and for the remaining 6 children with SSD this measure was the Bankson-Bernthal Test of Phonology Consonant Inventory (Bankson & Bernthal, 1990). All children with SSD scored within 1 *SD* of the mean or higher on a test of nonverbal intelligence (*M* = 109.29, *SD* = 14.96). Half the children with SSD presented with a language disorder and obtained a standard score below 87 (*M* = 76.2, *SD* = 9.52, range = 56-87) on the Structured Photographic Expressive Language Test-Preschool, Second Edition (SPELT-P2; Dawson et al., 2005), the cutoff reflecting good sensitivity and specificity for Developmental Language Disorder (DLD; Greenslade et al., 2009). The remaining 10 children with SSD did not present with concomitant language disorder and scored above this cutoff on the same measure (*M* = 102.4, *SD* = 9.23, range = 89-114). Children with TD did not have a history of speech-language assessment or intervention, and their parents did not express concerns regarding their child's communication development. They obtained scores within 1 SD of the mean or higher on the GFTA-3 (n = 20, *M* = 101.30, *SD* = 10.58, range = 87-121) and the

SPELT-P2 (*M* = 111.30; *SD* = 6.81, range = 96-123); see Figure 1 for the variation in SPELT-P2 scores between the groups.

** insert Figure 1 here **

Procedure

All children attended two testing sessions, each lasting approximately 60 minutes and separated by about one week. Children were assessed in a quiet room using the standardized measures described in the Participants section. Assessments were completed prior to the experimental task. Children were then brought into a single-walled sound booth, where they completed a looking-while-listening task in one of two conditions: (1) A Restrictor condition (R), during which movement of the tongue was restricted by a silicone fish on a lanyard commonly used for pediatric oral motor therapy (item OM8221, Therapy Shoppe; see Figure 2; this restrictor is similar to that used in Bruderer et al., 2015, and is the same as that used in Seidl et al., 2018); (2) A No Restrictor condition (NR), during which articulatory movement of the tongue was not restricted. The order of the testing conditions was counterbalanced within each group of participants (SSD, TD) such that half of the participants completed the experiment with the Restrictor (R) on day 1 and half completed the experiment with the Restrictor on day 2, and vice versa for the No Restrictor condition (NR) using a within-subjects design.

** insert Figure 2 here **

Participants could either sit alone in the booth, or on their caregiver’s lap, in which case the caregiver wore opaque sunglasses to obstruct their view of the visual stimuli. During the Restrictor session, children placed a cloth lanyard attached to the silicone fish around their necks and were instructed to place the tail/end portion of the silicone fish in their mouth (the fish’s tail in Figure 2), such that tongue movement to the alveolar and palatal places of articulation required for producing /s/ and /ʃ/ was restricted. All children complied with this instruction and kept the fish in their mouth throughout the task. Procedures on the No Restrictor (NR) day were identical except for the absence of the fish and lanyard.

Stimuli

Auditory stimuli for the looking-while-listening task were recorded by a female monolingual speaker of Midwestern American English. Eighteen target words, non-words, and foils (“shoe”/“sue”, “ship”/“sip”, “sheep”/“seep”, “sock”/“shock”, “soap”/“shope”, “sun”/ “shun”, “chair”, “fork”, “cup”, “apple”, “doll”, “book”) were recorded and spliced in Praat (Version 6.0.43; Boersma & Weenink, 2017) to follow the same token of “Look!” resulting in phrases such as “Look! Sheep!”.

Six yoked pairs of visual stimuli were created for use in the looking-while-listening task. There were a total of 18 trials (see Table 1) that were identical for both conditions (NR, R) and groups (SSD, TD): 6 using correctly pronounced words (CP; e.g., “shoe”, trial 8), 6 using mispronounced targets (MP; e.g., “sue”, trial 3), and 6 foils (F; e.g., “fork”, trial 11). This design meant that children saw each of those 6 yoked pairs of images (e.g., *fork* and *shoe*) repeated 3 times, but heard a different word each time depending on the type of trial: “fork” during F trial, “shoe” during CP trial, and “sue during MP trial. Foil trials served as a distraction from the main purpose of the experiment and enabled us to examine broad effects of the presence/absence of the restrictor. Images were equal in size, were presented side-by-side on a white background, and appeared an equal number of times across trials[2].

Each trial was 5000 msec long. This included a 1000 msec pre-naming period of silence followed by "Look" before the target word (presented 2000 msec into the trial). This was then followed by another period of silence until the end of the trial (at 5000 msec).

** insert Table 1 here **

Eye-Gaze Coding

The looking-while-listening task was video-recorded for offline coding of children's eye gaze. Videos were coded frame-by-frame using Supercoder (Hollich, 2005) by undergraduate research assistants (RAs) who were trained by the first author and who were blind to participant group. For each frame, RAs coded whether the child was looking left, right, or away from the screen.

Reliability

We randomly chose 8 videos (i.e., 10%) from our sample of 80 videos (2 for each of the 40 participants) to examine the reliability of eye-gaze coding. These files were coded separately by two RAs who were also trained by the first author and were performing similar coding of data collected from other projects. For each of the 144 trials in the reliability sample, we calculated a difference score for the number of frames coded as looks to the right between the two coders (M = .055; SD = 9.67; range = -86 – 32; 95% CI [-1.53 – 1.64]) and a difference score for the number

---

[2] Despite controlling for the frequency with which images were displayed, the current design meant that we asked about the visual target in CP and MP trials (e.g., "soap", "shope") twice as frequently as we asked for the target in F trials (e.g., "book"). We address this issue in our analyses of eye-gaze data.

of frames coded as looks to the left (*M* = .145; *SD* = 5.76; range = -26 – 31; 95% CI [-.803 – .145]). Most of the difference scores clustered around one frame difference between coders, and matched pairs t-tests showed that coders did not differ in the number of frames they coded as looks to the right (*t*(143) = .0689, *p* = .945) or to the left (*t*(143) = .303, *p* = .761) in each trial. Note that RAs who performed the coding and those who re-coded a portion of the data for reliability went through the same rigorous training process by the first author, during which they were required to achieve equal levels of reliability to those reported above before they could code data from any research project.

Results

Before conducting any analyses, we took a few data cleaning steps that were driven by our experimental design. First, given that our design focuses on children's ability to perceive the targets of words that include /s/ and /ʃ/, we removed any trials that could interfere with children's performance due to the presence of an object whose label could begin with an affricate. Specifically, we removed trials number 9 and 15 during which children saw pictures of *chair* and *sun*, and heard the targets "shun" and "sun" while they also saw the non-target *chair*. These were removed out of a concern that the presence of the non-target, "chair", whose label begins with a palatal affricate which shares a place of articulation with the target sound /ʃ/, in "shun", might impact children's performance in a way that is distinct from other CP and MP trials in which the labels of distractor objects do not begin with such affricates (e.g., "cup"). Next, given that foil trials were added to distract children from the main purpose of the experiment, and to allow us to examine potential broad speech perception effects of the restrictor, we analyze data from these trials separately.[3]

Analyzing Eye-Gaze Data

For each participant, we calculated the proportion of looking at the target and the proportion of looking at the distractor in each frame (i.e., 1/30 of a second), for each of the trial types (CP, MP) averaged across trials (5 trials per type). We then used these measures to calculate the log-gaze proportion ratio, calculated as the log of the proportion of fixations to the target divided by the proportion of fixations to the distractor (i.e., log-gaze proportion ratio = log(proportion fixations to target / proportion fixations to distractor)). This measure represents children's relative preference for fixating on the target over the distractor and allows for a clear visualization of the data. Zero values represent equivalent fixation to the target and distractor, whereas values less than 0 indicate a preference for the distractor, and values greater than zero indicate a preference for the target.

Figure 3 shows children's relative preference for the target over the distractor throughout the *entire trial*. As is evident in the figure, children in both diagnosis groups (SSD, TD) across the two conditions (NR, R) show a clear preference for the target throughout the entire trial (i.e., most of the log gaze is above zero), with the magnitude of this preference increasing after the onset of

---

[3] To test the specificity of the effect of the restrictor, it would have been *ideal* to model foil trials along with CP and MP trials –i.e., include a trial-type factor with 3 levels– but our experimental design (see footnote 1, and corresponding text) did not allow us to treat these 3 types of trials equally and weight them similarly in a statistical model. That is because CP and MP trials were advantaged by sharing the same visual targets. This issue may limit the conclusions that can be drawn about the specificity of the effect of the restrictor. Future studies with balanced numbers of Foil and CP/MP trials will allow for a more rigorous test of the specificity of the effect of the restrictor.

the target word (at 2000 msec), and changing across trial types, conditions, and diagnosis groups. This pattern may stem from the fact that the target of CP and MP trials (e.g., “shoe” and “sue”) was represented by the same picture twice (i.e., *shoe*), which may have triggered children to anticipate that we will ask about those CP and MP targets (e.g., “shoe”). In order to explore this further, we visually inspected the effect of trial order focusing on the pre-naming period (i.e., the first 2000 msec of each trial; see Figure 1 in the Supplementary Materials), and found that children in both groups across the two conditions (on either day) showed a clear preference for looking at the target in most CP and MP trials.

** insert Figure 3 here **

In the analyses we report on next, which explore fine-grained differences in gaze patterns in the post-naming period across the diagnostic groups (SSD, TD), the two conditions (NR, R), and trial types (CP, MP), we correct for the above-mentioned preference for targets by utilizing a new dependent measure. Below, we detail the steps we took to calculate this measure, all of which were informed by our results, as seen in Figure 3, and by previous literature (e.g., Law & Edwards, 2015; Marchman & Fernald, 2008; Reuter et al., 2019).

First, we refine our window of analysis to a post-naming window between 300 and 1800 msec from the onset of the target word, resulting in a window of 1500 msec (Law & Edwards, 2015; Marchman & Fernald, 2008). We also use a pre-naming time-window, which starts at the beginning of the trial and ends 300 msec from the onset of the target word (which occurs at 2000 msec), resulting in a window of 2300 msec.

Second, to solve any potential biases in gaze patterns stemming from the preference for the targets during the pre-naming period, we offset the proportion of fixations to the target by subtracting the proportion of looking at the target in the pre-naming window from the proportion of looking at the target in the post-naming window. Hence, the dependent measure we use in the remainder of this manuscript represents the ***increase in the proportion of fixations to the target*** in the post-naming period compared with the pre-naming baseline.

Prior to performing statistical analyses, we performed some extra steps to clean the dataset. Specifically, we removed all trials in which children looked at the screen for less than 20% of the duration of either the pre-naming or post-naming window of analysis or both. Out of a total of 800 trials (after removing foil trials and "chair" trials) across all participants and conditions, we removed 61 trials (less than 10% of our sample, i.e., 7.6%) with less than 20% looking either during the pre-naming window, post-naming window or both.

Statistical Models

To reiterate, given our interest in examining the effect of the restrictor on children's perception and recognition of CP and MP target words that include /s/ and /ʃ/, and whether such an effect is *specific* to the perception of those sounds or if it interferes with perceiving and recognizing targets that do not include /s/ and /ʃ/, as in foil trials, we report on separate statistical models: first for CP and MP trials and next for foil trials only. Alpha level for all main and interaction effects was .05, and we used the Bonferroni correction where necessary to correct for multiple comparisons.

CP and MP trials

As a first step, we examined whether the day in which children were tested with (R) or without the restrictor (NR) had any impact on their performance. Given that we counterbalanced this factor across our sample (half of the children in each diagnosis group, i.e., n=10, were tested with the restrictor (R) on day 1 whereas the other half were tested with the restrictor (R) on day

2), we did not expect *day* to have an impact on the increase in proportion of fixations towards the target. We fitted a mixed linear regression model with *diagnosis* (SSD, TD), *condition* (NR, R), *trial type* (CP, MP), and *day* (1, 2) as fixed effects, along with all 2-way, 3-way interactions, and the 4-way interaction (*diagnosis* x *condition* x *trial type* x *day*), and *subject* as a random effect. The model revealed a main effect of *trial type* (Estimate = .034; SE = .013; 95% CI [.009 - .06]; *F*(1,689.9) = 7.02, *p* = .0082), and a significant *diagnosis* by *condition* interaction (Estimate = -.03; SE = .013; 95% CI [-.055 - -.004]; *F*(1,689.9) = 5.22, *p* = .0226) along with a significant *diagnosis* by *day* interaction (Estimate = .034; SE = .013; 95% CI [.0082 - .056]; *F*(1,689.9) = 6.726, *p* = .0097) on the increase in proportion of fixations to the target; all other fixed effects and their interactions were not significant (*p*s > .15).

To explore what drives the *diagnosis* by *day* interaction, we ran post-hoc pairwise comparisons. Our main interest here was to explore whether the day of testing (regardless of condition, i.e., NR, R) had a different impact on the increase in proportion of fixations to the target across the two diagnoses. Hence, we only report the comparisons of day of testing within the diagnosis groups with a Bonferroni corrected alpha of .025 for multiple comparisons. Children with SSD showed a larger increase in their looking to the target on the first day of testing (*M* = .194; *SD* = .37; 95% CI [.14 - .248]) compared to the second day of testing (*M* = .123; *SD* = .406; 95% CI [.063 - .183]); however, that difference was not statistically significant (*t* = 1.88; *p* = .0602). Children with TD on the other hand showed a larger increase in their looking to the target on the second day of testing (*M* = .172; *SD* = .378; 95% CI [.117 - .226]) compared to the first day of testing (*M* = .105; *SD* = .301; 95% CI [.062 - .148); yet again, this difference was not statistically significant (*t* = -1.78; *p* = .0747). Given that we do not have evidence for a statistically significant difference in the increase in proportion of fixations towards the target between the two diagnostic groups across the two days of testing, we removed the fixed effect *day* and all interaction effects that include *day* from our subsequent analyses.

Next, we re-fitted a mixed linear regression model with *diagnosis* (SSD, TD), *condition* (NR, R), and *trial type* (CP, MP) as fixed effects, along with all 2-way and 3-way interactions, and *subject* as a random effect. The model revealed a main effect of *trial type* (Estimate = .035; SE = .013; 95% CI [.009 - .06]; $F(1,696.2) = 7.188$, $p = .0075$) and a significant *diagnosis* by *condition* interaction (Estimate = -.03; SE = .013; 95% CI [-.056 - -.004]; $F(1,695.9) = 5.45$, $p = .0198$) on the increase in proportion of fixations to the target; all other fixed effects and their interactions were not significant (*p*s > .17)[4]. The significant *trial type* effect is evident in a larger increase in proportion of fixations to the target in CP trials ($M = .184$; $SD = .36$; 95% CI [.147 - .221]) compared to MP trials ($M = .112$; $SD = .37$; 95% CI [.074 - .15]); see Figure 4.

** insert Figure 4 here **

Last, we used post-hoc pairwise comparisons to explore the *diagnosis* by *condition* interaction effect in more detail, with a Bonferroni-corrected alpha of .025. Here too, we were interested in the between-condition comparisons across each of the diagnoses separately (SSD: NR vs. R; TD: NR vs. R). The pairwise t-tests provide evidence that children with SSD showed a smaller increase in the proportion of fixations towards the target (in CP and MP trials) in the R condition ($M = .115$; $SD = .409$; 95% CI [.055 - .175]) compared to the NR condition ($M = .204$; $SD = .362$; 95% CI [.151 - .258]; $t = -2.35$; $p = .0192$); whereas a similar comparison in the TD

---

[4] We performed the same statistical analyses *with* trials 9 and 15 that use the target image *chair* and the results do not change. Specifically, we fitted a mixed linear regression model with *diagnosis* (SSD, TD), *condition* (NR, R), *trial type* (CP, MP) as fixed effects, along with all 2-way and 3-way interactions, and *subject* as a random effect. The model revealed a main effect of *trial type* ($F(1,840.6) = 8.66$, $p = .0033$), as well as a significant *diagnosis by condition* interaction ($F(1,840.7) = 7.75$, $p = .0055$) on the increase in proportion of fixations to the target; all other fixed effects and their interactions were not significant (*p*s > .2).

group was not statistically significant (R: *M* = .155; *SD* = .33; 95% CI [.108 - .202]; NR: *M* = .121; *SD* = .355; 95% CI [.07 - .172]; *t* = .94; *p* = .347); see Figure 5. Note that the difference between TD and SSD children in their performance *without* the restrictor is not significant (TD: *M* = .121; *SD* = .355; 95% CI [.07 - .172]; SSD: *M* = .204; *SD* = .362; 95% CI [.151 - .258]; *t* = 1.82; *p* = .071).

** insert Figure 5 here **

These results show that children with SSD were more influenced by the restrictor than TD children. The impact of the restrictor did not depend on the type of trial (CP, MP) and children with SSD performed better without the restrictor than with it. Crucially, despite the wide variability in SPELT-P2 scores among children with SSD (see Figure 1), we found no evidence for any significant linear correlation between the increase in the proportion of children's fixations to the target and their SPELT-P2 scores in either condition (*p*s > .6); SPELT-P2 scores were also not correlated with the performance of children with TD (*p*s > .6). These results show that the performance in our speech perception task was not related with children's language skills.

Foil trials

We fitted a mixed linear regression model only for foil trials with *diagnosis* (SSD, TD) and *condition* (NR, R) as fixed effects, along with the 2-way *diagnosis* by *condition* interaction, and *subject* as a random effect; unlike the models reported above for CP and MP trials, none of the effects in this model were significant (*diagnosis*: Estimate = -.0414; SE = .024; 95% CI [-.09 - .007]; *F*(1,38.6) = 2.967, *p* = .093; *condition*: Estimate = .0035; SE = .0192; 95% CI [-.034 - .041]; *F*(1,333.5) = .034, *p* = .853; *diagnosis x condition*: Estimate = -.0023; SE = .0192; 95% CI [-.04 - .035]; *F*(1,333.5) = .015, *p* = .901), suggesting that there is no evidence that children in the two diagnostic groups across the two conditions treated the foil trials differently; see Figure 6. This result is in line with developmental work (Bruderer et al., 2015; Choi et al., 2015) showing that restrictions to the movement of articulators exert a specific effect on speech perception rather than a broad one (though see footnote 3).

** insert Figure 6 here **

Discussion

In the current study we examined whether temporarily restricting articulatory movements affects the speech perception abilities of 48- to 74-month-old children with TD and SSD. In a looking-while-listening task, children were presented with two visual images (e.g., of a *shoe* and a *fork*) and heard “Look! Fork!” in Foil trials (F), “Look! Shoe” in Correctly-pronounced (CP) trials, or “Look! Sue!” in Mispronounced trials (MP). Children were tested twice: on one day with a restriction to their articulators and on another day without this restriction. Analyses of their fixation patterns during the experiment showed that, as expected, children in both groups across the two conditions and for all three trial types increased their fixations to the target shortly after the onset of the target word; hence, they were able to perceive the target words regardless of pronunciation and restrictor and to accurately orient to the target picture upon hearing the target word. However, there was a robust condition by diagnosis effect that revealed that, unlike TD children who performed equally across the two experimental conditions (restrictor, no-restrictor), children with SSD showed a greater increase in their fixations to the target in the condition *without* restriction to their articulators, compared to the condition when such restrictions were in place.

Bruderer and colleagues (2015) reported that 6-month-olds are sensitive to the presence of a specific articulatory restrictor during speech perception of a non-native contrast. These results provide supporting evidence that sensorimotor cues may affect perception. In this infant work, a case was made for the importance of assessing learners with impaired speech production: “A hypothesis that follows from the findings is that impairment to the articulatory system could be a hindrance to speech perception and language development” (Bruderer et al., 2015, pg. 13535). Testing this hypothesis was a central aim of the present work. Children with SSD are characterized by their deficits in production (e.g., Gierut, 1998; Shriberg et al., 2010), though perceptual deficits are frequently implicated (e.g., Brosseau-Lapré et al., 2020). Thus, it was predicted that children with SSD would show weaker perceptual abilities than their TD peers, even in the condition in which no restrictor to would-be articulators was included. Further, we expected there might be developmental continuity from the infant work, with TD children showing sensitivity to the presence of the restrictor (though perhaps less than infants), while children with SSD, with their impaired sensorimotor mappings, would not show such sensitivity.

Neither of our initial predictions were supported. First, children with SSD showed similar perceptual abilities to their TD peers when no restrictor was present. In short, perception of the words included in this study was not noticeably weaker in children with SSD than those with TD. The second prediction concerned the influence of the articulatory restrictor on perception. In this case, again counter to our predictions, the children with TD showed no effect; the restrictor did not change their perceptual abilities in a measurable way. However, the children with SSD *were* sensitive to the presence of the restrictor: there was a decrement in performance in their speech perception in the looking-while-listening task when the articulatory restrictor was in place.

In the remainder of the discussion, we turn to potential accounts of why the presence of an articulatory restrictor affected the speech perception of children with SSD but not their TD peers. In short, we found that 4- to 6-year-old children with SSD are similar to infants (Bruderer et al., 2015) in that they also show sensitivity to a restrictor to would-be articulators during perception of target native sounds, while slightly younger children with TD do not show this effect and are robust to the perturbation we tested. We interpret this both developmentally and clinically. That SSD children perform like infants is not surprising; the surprising component of the results of this work is that, in typical learners, effects of the disruptor seem to disappear by the late preschool years. This developmental shift was unexpected and merits additional investigation since it is unclear at which age this robustness might emerge and/or whether it generalizes to other sound contrasts.

Another possible explanation for our findings is that children with SSD do not show sensorimotor mapping deficits or delays, but rather attend more poorly to the speech stimuli either with or without a restrictor in place, or perhaps, are globally perturbed by a sensory restrictor. We think that attentional factors do not explain our results for a few key reasons. First, the performance of children with SSD was similar to that of their TD peers when the restrictor was not in place. Thus, there is no reason to believe that these children were not attending to the task. Second, and even more central to a sensorimotor mapping disruption hypothesis, children with SSD only showed reduced sensitivity in perception with the restrictor when the implicated sounds (/s/ and /ʃ/) incorporated the place of articulation impeded by the restrictor. Perception of the foil words, which included a preponderance of labial and velar consonants as well as vowels in word onset position, were not affected by the restrictor (which in this case, did not directly impact would-be articulators). Thus, there is evidence for specific sensorimotor effects; an attentional account will not suffice to explain these results. That said, a potential alternative explanation, and one that needs to be addressed in future work, is that /s/ and /ʃ/ may be more perceptually vulnerable because they are later developing than the sounds incorporated in the foils and that such effects might not have been seen in earlier developing sounds. However, the present results provide preliminary evidence that it is not the global articulatory perturbation, but rather the specific blocking of sensorimotor feedback that is critical for the production of /s/ and /ʃ/ that influences the perception of these sounds in children with SSD.

Thus far, a fully developmental account that incorporates phonetically specified sensorimotor mapping seems most plausible. However, when turning to comparing the present results to those obtained from the production task conducted by Seidl et al. (2018), it seems that a deficit in sensorimotor mapping may be implicated in children with SSD. In this previous work, the aim was to determine if the presence of a restrictor during perception influenced a production after the restrictor was removed. Results showed that in the production domain, unlike the perception domain tested in the present work, children with SSD showed *no* measurable sensitivity to the presence of a restrictor, while those with TD *did* show this sensitivity. Specifically, Seidl and colleagues (2018) focused on how children produce contrastive acoustic features for /s/ and /ʃ/ in a novel word repetition task; before the novel word repetition task, children were exposed to the novel words with /s/, /ʃ/ onsets along with their referents (i.e., novel creatures) either with or without an articulatory restrictor. Though both SSD and TD groups produced distinct /s/ and /ʃ/ categories, different response patterns were observed across groups. TD children showed sensitivity to the restrictor (through hyper articulation in production) while SSD children showed no differences in production whether the articulatory restrictor was present or not during the novel word learning period. In contrast, the perception results obtained in the present work reveal that only children with SSD, but not their TD peers, are sensitive to the restriction of sensorimotor feedback. It is important to consider how these results may be reconciled in line with DIVA and other accounts of sensorimotor linkages.

A simple explanation is that differing effects are driven by task complexity. Thus, pre-schoolers with SSD (like infants) are sensitive to links across sensorimotor and auditory components only in very simple perceptual and temporally constrained tasks (like our perception task), but insensitive in more demanding production tasks (like the one in Seidl et al., 2018). Why do TD pre-schoolers not show an effect of the restrictor in simple perceptual tasks? Recall that we started this paper with examples of how talkers must perceive and produce intelligible speech while eating a sandwich, sucking a pacifier or a pen, etc. We suggest that, by the time TD children reach preschool, their sensorimotor systems respond rapidly and resiliently to online perturbations to would-be articulators (such as having a toy/food in their mouth) during speech perception. Thus, in our task, TD preschool learners have already adapted to perturbations of this sort, but infants (results in Bruderer et al., 2015) and children with SSD do not show such flexibility given their feed-forward systems are still developing (infants) or are engaged in protracted development given the nature of their disorder (SSD).

While the simple perception task may be explained as above, with reference to development, we still need to reconcile the results of the more complex production task in Seidl and colleagues (2018) with ours to explain why the pattern of results is different in the more complex task. Recall that, in the production task in Seidl et al., only children with TD (not SSD) showed sensitivity to sensorimotor disruption to would-be articulators during perception on productions that followed (i.e., they hyperarticulated the contrast after the restrictor was removed). From these findings, children with SSD were presumed to show an impairment in sensorimotor feedback processes. It was suggested that lack of sensitivity to the restrictor was consistent with other perceptual and production deficits attested in children with SSD and that observed perceptual and production deficits may occur due to an impairment in linking somatosensory information to vocal motor schema. The present results are quite distinct; in a purely perceptual task, only children with SSD were sensitive to the presence of an articulatory restrictor. We argue that the production task differs in two potentially important ways from the one conducted in the present perceptual study, again, related to task complexity. First, to show hyper articulation, the influence of the restrictor must be held briefly in memory since the child removes the restrictor and *then* produces the novel word. Second, children with SSD have known deficits in novel word repetition tasks (e.g., Shriberg et al., 2009; Vuolo & Goffman, 2020). In sum, both the memory component and the novel word repetition component of the task could contribute to SSD children’s lack of sensitivity to the presence of the restrictor. While these differing patterns of results from these two tasks on perception and production present more questions than they answer, these data suggest that perception-production loops play differential roles in both TD and SSD learners’ task performance.

In addition to introducing new questions to be investigated, the different findings from these two distinct, yet related, tasks point out some limitations of the DIVA model as applied to developmental phenomena. DIVA is a model of speech motor control that incorporates somatosensory, auditory, and articulatory interactions (e.g., Tourville & Guenther, 2011). In speech production, as pointed out by Levelt and colleagues (1999), these speech motor levels must interface with higher linguistic processes, such as the lexicon. The earlier results from preverbal and even pre-canonical infants make clear that production experience and skill are not obligatory for sensitivity to somatosensory input to occur. In preschool aged children who are TD, a sensorimotor disruption is observed in a production task that involves mapping a novel referent to a novel word form. This task increases memory load via a delay between listening (with the restrictor in place) and talking (with the restrictor removed). When only perception is involved, unlike infants, children who are TD show no such disruption. Children with SSD, however, behave similarly to infants in a perception only task that incorporates known words. That is, as with infants, articulatory restrictions influence speech perception in children with SSD. In summary, somatosensory, auditory, and articulatory loops that form the basis of DIVA require increased nuance, with differential relationships and weightings emerging via developmental, linguistic, and cognitive factors.

To conclude, the present results reveal an unexpected developmental story and provide insight into the nature of SSD. In typical development, it appears that the motor system influences infants' speech perception in simple tasks (Bruderer et al., 2015) and, based on the present findings, the same is true for children with SSD. However, at least some of these motor influences on perception disappear in TD preschool aged learners during simple tasks – thus explaining how, as adults, we can perceive speech well despite sensorimotor restrictions (e.g., eating a sandwich) during our listening to speech. SSD is characterized as a production impairment. Based on previous work, we suggest that children with SSD may have a representational model that integrates sensorimotor information but may show deficits in mapping that model to production processes which may contribute to their production difficulties. The present work supports that production and sensorimotor mapping are heavily implicated in SSD since disruption in the sensorimotor system has clear results for children's speech perception. Yet, counter to predictions from Bruderer et al. (2015), we did not observe a clear relationship between language deficit (as observed in the SPELT task performance) and performance on our task.

While the present results provide a first report of how somatosensory input influences speech perception in children with and without SSD, future work will be needed to 1) replicate our findings with larger samples, more contrasts, and different somatosensory disruptions, 2) test the specificity of the effect of the restrictor in trials that are equally weighted, 3) test the developmental effects we observed under various experimental conditions (e.g., varying task complexity and outcome measures), and 4) examine whether the proposed impact of somatosensory disruptions on speech processing is generalizable and influential in real life listening conditions. Crucially, and in order to reconcile the infant and adult literature with our current findings, more work is needed to examine the *developmental trajectory* of the linkages between speech perception and somatosensory mappings.

Acknowledgements

The authors declare no competing interests, or conflicts of interest regarding the funding of the study. We thank Katie Isbell for her help with testing children, the numerous undergraduate research assistants who coded the looking-while-listening data and the families who participated in the study. This work was supported by NIH/NIDCD R01DC018410.

Data Availability Statement

All visual and auditory stimuli, as well as the raw data spreadsheets are available on the project osf page: https://osf.io/4dbx5/?view_only=67595bf0e7ba4b0fbe13149a31b09ab7

References

Bankson, N. W., & Bernthal, J. E. (1990). *Bankson-Bernthal Test of Phonology*. Riverside Press.

Boersma, P. & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 02 January 2017 from https://www.praat.org.

Brosseau-Lapré, F., Schumaker, J., & Kluender, K. R. (2020). Perception of medial consonants by preschoolers with and without speech sound disorders. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3600-3610. https:/doi.org/10.1044/2020_JSLHR-20-00146

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, *112*(44), 13531-13536. https:/doi.org/10.1073/pnas.1508631112

Choi, D., Bruderer, A. G., & Werker, J. F. (2019). Sensorimotor influences on speech perception in pre-babbling infants: Replication and extension of Bruderer et al. (2015). *Psychonomic Bulletin & Review*, *26*(4), 1388-1399. https://doi.org/10.3758/s13423-019-01601-0

Choi, D., Dehaene-Lambertz, G., Peña, M., & Werker, J. F. (2021). Neural indicators of articulator-specific sensorimotor influences on infant speech perception. *Proceedings of the National Academy of Sciences*, *118*(20). https:/doi.org/10.1073/pnas.2025043118

Dawson, J., Stout, C., Eyer, J., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test – Preschool, Second Edition*. Janelle Publications.

Edwards. J. (1992). Compensatory speech motor abilities in normal and phonologically disordered children. *Journal of Phonetics*, *20*, 189-207. https:/doi.org/10.1016/S0095-4470(19)30622-9

Edwards, J., Fox, R. A., & Rogers, C. L. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy. *Journal of*

Speech, Language, and Hearing Research, 45, 231-242. https:/doi.org/10.1044/1092-4388(2002/018)

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361-377. https:/doi.org/10.3758/bf03193857

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature, 462*(7272), 502-504. https:/doi.org/10.1038/nature08572

Gierut, J. A. (1998). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research, 41*(1), S85-S100.

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation, Third Edition*. Pearson Assessments.

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Phonographic Expressive Language Test-Preschool: Second Edition. *Language, Speech, and Hearing Services in Schools, 40*(2), 150-160. https:/doi.org/10.1044/0161-1461(2008/07-0049)

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350–365. https:/doi.org/10.1016/j.jcomdis.2006.06.013

Guenther, F. H. (2016). Neural Control of Speech. MIT Press, Cambridge, MA.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics, 25*(5), 408–422. https:/doi.org/10.1016/j.jneuroling.2009.08.006

Hearnshaw, S., Baker, E., & Munro, N. (2018). The speech perception skills of children with and without speech sound disorder. *Journal of Communication Disorders*, *71*, 61-71. https:/doi.org/10.1016/j.jcomdis.2017.12.004

Hearnshaw, S., Baker, E., & Munro, N. (2019). Speech perception skills of children with speech sound disorders: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, *62*(10), 3771-3789. https://doi.org/10.1044/2019_JSLHR-S-18-0519

Hollich, G. (2005). *Supercoder: A program for coding preferential looking*. [Computer Software]. West Lafayette, IN: Purdue University.

International Expert Panel on Multilingual Children's Speech (2012). Multilingual children with speech sound disorders: Position paper. Research Institute for Professional Practice, Learning and Education (RIPPLE), Charles Sturt University. Bathurst, NSW, Australia Retrieved from http://www.csu.edu.au/research/multilingualspeech/position-paper.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences*, *106*(4), 1245-1248.

Law, F., & Edwards, J. R. (2015). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development*, *11*(4), 331-355. https:/doi.org//10.1080/15475441.2014.961066

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

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431–461.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36. https://doi.org/10.1016/0010-0277(85)90021-6

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental science*, *11*(3), F9-F16. https:/doi.org/10.1111/j.1467-7687.2008.00671.x

Masapollo, M., & Guenther, F. H. (2019). Engaging the articulators enhances perception of concordant visible speech movements. *Journal of Speech, Language, and Hearing Research*, *62*(10), 3679-3688. https://doi.org/10.1044/2019_JSLHR-S-19-0167

McAllister Byun, T. (2012). Bidirectional perception–production relations in phonological development: evidence from positional neutralization. *Clinical Linguistics & Phonetics*, *26*(5), 397-413. https:/doi.org/10.3109/02699206.2011.641060

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children’s word learning. *Developmental Psychology*, *55*(8), 1656–1665. https:/doi.org/10.1037/dev0000754

Roepke, E., & Brosseau-Lapré, F. (2019). Perception of sibilants by preschool children with overt and covert sound contrasts. *Journal of Speech, Language, and Hearing Research*, *62*(10), 3763-3770. https:/doi.org/10.1044/2019_JSLHR-S-19-0127

Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*, *54*(2), 193-208. https://doi.org/10.1044/jshd.5402.193

Sénéchal, M., Ouellette, G., & Young, L. (2004). Testing the concurrent and predictive relations among articulation accuracy, speech perception, and phoneme awareness. *Journal of Experimental Child Psychology*, *89*(3), 242-269. https:/doi.org/10.1016/j.jecp.2004.07.005

Seidl, A., Brosseau-Lapré, F., & Goffman, L. (2018). The impact of brief restriction to articulation on children's subsequent speech production. *The Journal of the Acoustical Society of America*, *143*(2), 858-863. https:/doi.org/10.1121/1.5021710

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The Syllable Repetition Task (SRT). *Journal of Speech, Language, and Hearing Research*, *52*(5), 1189–1212. https://doi.org/10.1044/1092-4388(2009/08-0047)

Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeny, J. L., ... & Wilson, D. L. (2010). Extensions to the speech disorders classification system (SDCS). *Clinical Linguistics & Phonetics*, *24*(10), 795-824. https:/doi.org/10.3109/02699206.2010.503006

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147-166. https:/doi.org/10.1016/S0010-0277(00)00081-0

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2014a). Auditory–motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders*, *47*, 17-33. https:/doi.org/10.1016/j.jcomdis.2014.01.001

Terband, H., Van Brenk, F., & van Doornik-van der Zee, A. (2014b). Auditory feedback perturbation in children with developmental speech sound disorders. *Journal of Communication Disorders*, *51,* 64-77. https://doi.org/10.1016/j.jcomdis.2014.06.009

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952-981. https://doi.org/10.1080/01690960903498424

Towne, R. (1994). Effects of mandibular stabilization on the diadochokinetic performance of children with phonological disorders. *Journal of Phonetics, 22*(3), 317–332.

Vuolo, J., & Goffman, L. (2020). Vowel accuracy and segmental variability differentiate children with developmental language disorder in nonword repetition. *Journal of Speech, Language, and Hearing Research, 63*(12), 3945-3960. https:/doi.org/10.1044/2020_JSLHR-20-00166

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49-63. https://doi.org/10.1016/S0163-6383(84)80022-3

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701-702. https:/doi.org/10.1038/nn1263

Yeung, H. H., & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science*, *24*(5), 603-612. https://doi.org/10.1177/0956797612458802

Figure and Table Titles and Legends

Figure 1

Violin Plot of SPELT Scores for Children with SSD and Children with TD

Figure 2

The Silicone Fish Used to Restrict Tongue Movement in the Restrictor (R) condition

Table 1

Details of all 18 Trials in the Looking-While-Listening Task (Attention Getters Appeared Between Trials)

Figure 3

The Log-Gaze Proportion Ratio During the Entire Trial by Trial Type (CP, MP) Across the two Conditions (No Restrictor (NR), Restrictor (R)) and two Diagnosis Groups (SSD, TD).

Note. Zero values represent equivalent fixations to the target and distractor, values above zero represent preference for the target, and values below zero represent preference for the distractor. The dashed line at 2000 msec represents the onset of the target word. The shaded areas around the lines represent the bootstrap confidence region of the fit for 95% of the smoother fits.

Figure 4

A Violin Plot of The Increase in Proportion of Fixations to the Target by Trial Type

Figure 5

The Increase in Proportion of Fixations to the Target During CP and MP Trials (combined) in Each Condition (NR, R) for Both Diagnostic Group (SSD, TD). Error bars represent standard error.

Figure 6

The Increase in Proportion of Fixations to the Target During Foil Trials in Each Condition (NR, R) for Both Diagnostic Group (SSD, TD). Error bars represent standard error.

Supplementary Materials

Figure 1 in the Supplementary Materials shows the proportion of fixations to the Target and Distractor during the *pre-naming* phase of CP and MP trials in each of the two conditions across the two days and two diagnostic groups. This figure shows that children in both groups across the two conditions (on either day) have a clear preference for looking at the target in most CP and MP trials. We corrected for this preference by reporting on the *increase* in the proportion of looking at the target from this pre-naming period.