ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Auditory and Somatosensory Development for Speech in Later Childhood

Samantha A. Ayala,[a] Amanda Eads,[a] Heather Kabakoff,[b] Michelle T. Swartz,[c] Douglas M. Shiller,[d] Jennifer Hill,[e] Elaine R. Hitchcock,[f] Jonathan L. Preston,[g] and Tara McAllister[a]

[a] Department of Communicative Sciences and Disorders, New York University, NY [b] Department of Neurology, New York University Grossman School of Medicine, NY [c] Department of Speech-Language Pathology, Thomas Jefferson University, Philadelphia, PA [d] École d'orthophonie et d'audiologie, Faculté de medicine, Université de Montréal, Québec, Canada [e] Center for Practice and Research at the Intersection of Information, Society, and Methodology, New York University, NY [f] Department of Communication Sciences and Disorders, Montclair State University, NJ [g] Department of Communication Sciences and Disorders, Syracuse University, NY

ARTICLE INFO

Article History:
Received August 25, 2022
Revision received November 29, 2022
Accepted December 30, 2022

Editor-in-Chief: Cara E. Stepp
Editor: Raymond D. Kent

https://doi.org/10.1044/2022_JSLHR-22-00496

ABSTRACT

Purpose: This study collected measures of auditory-perceptual and oral somatosensory acuity in typically developing children and adolescents aged 9–15 years. We aimed to establish reference data that can be used as a point of comparison for individuals with residual speech sound disorder (RSSD), especially for RSSD affecting American English rhotics. We examined concurrent validity between tasks and hypothesized that performance on at least some tasks would show a significant association with age, reflecting ongoing refinement of sensory function in later childhood. We also tested for an inverse relationship between performance on auditory and somatosensory tasks, which would support the hypothesis of a trade-off between sensory domains.
Method: Ninety-eight children completed three auditory-perceptual tasks (identification and discrimination of stimuli from a "rake"–"wake" continuum and category goodness judgment for naturally produced words containing rhotics) and three oral somatosensory tasks (bite block with auditory masking, oral stereognosis, and articulatory awareness, which involved explicit judgments of relative tongue position for different speech sounds). Pairwise associations were examined between tasks within each domain and between task performance and age. Composite measures of auditory-perceptual and somatosensory functions were used to investigate the possibility of a sensory trade-off.
Results: Statistically significant associations were observed between the identification and discrimination tasks and the bite block and articulatory awareness tasks. In addition, significant associations with age were found for the category goodness and bite block tasks. There was no statistically significant evidence of a trade-off between auditory-perceptual and somatosensory domains.
Conclusions: This study provided a multidimensional characterization of speech-related sensory function in older children/adolescents. Complete materials to administer all experimental tasks have been shared, along with measures of central tendency and dispersion for scores in two subgroups of age. Ultimately, we hope to apply this information to make customized treatment recommendations for children with RSSD based on sensory profiles.

Existing theoretical frameworks of speech sound production (e.g., Guenther, 2016; Houde & Nagarajan, 2011; Parrell & Houde, 2019) identify auditory and somatosensory targets and feedback as central components in the control of speech. These models describe stored representations for speech sounds as learned associations between motor commands and sensory outcomes in both auditory-perceptual and somatosensory domains. Auditory targets are characterized as time-varying regions in a multidimensional auditory–acoustic space; they are initially formed by tracking the statistics of speech input from other individuals in the learner's environment. As the learner begins to

Correspondence to Tara McAllister: tkm214@nyu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

produce their own speech output, they can compare the auditory feedback they receive from these attempts with the external targets and begin to refine their stored motor commands to hit these targets more accurately. At the same time, the learner experiences the sense of touch, pressure, and movement associated with various configurations of the vocal tract, and they begin to lay down somatosensory targets associated with different speech sounds. Mature speakers possess stored motor commands (feedforward plans) that can be executed with limited reliance on auditory and somatosensory feedback (Guenther, 2016). However, there is ongoing monitoring of the speech signal by the feedback control system, which compares the sensory consequences of execution of the feedforward plan to the system's stored auditory and somatosensory targets and sends corrective commands as needed. Stored representations for speech sounds are updated throughout the life span as the vocal tract changes and speakers refine their sensory targets based on accumulated experience.

There is abundant evidence that both auditory function and somatosensory function play an active role in the development of speech production, as well as its ongoing control in mature speakers. Previous literature suggests that differences in the precision and consistency of speech sound production can be linked to individual differences in auditory-perceptual function (e.g., Newman, 2003; Perkell, Guenther, et al., 2004; Perkell, Matthies, et al., 2004; Villacorta et al., 2007) and somatosensory function (e.g., Ghosh et al., 2010; Wohlert & Smith, 1998). In both auditory and somatosensory domains, it is predicted that speakers with more narrowly specified sensory targets for speech will also produce speech sounds with greater precision than those with broader targets. With regard to auditory-perceptual function, theoretical models suggest that individuals with greater acuity are more sensitive to fine-grained detail in the acoustic signal, leading them to learn more tightly specified goal regions that are also further apart from neighboring phonemes (Perkell, 2012). Similarly, speakers with high somatosensory acuity are predicted to form tighter somatosensory goal regions, which could result in behaviors such as more consistent contact of the articulators on a given region in the vocal tract, for example, tongue tip to alveolar ridge (Ghosh et al., 2010; Perkell, 2012).

Sensory Acuity, Speech Sound Disorder, and Personalized Learning

This article examined speech-related sensory function in children and adolescents aged 9–15 years, with the goal of collecting reference data that can be used to characterize auditory-perceptual and somatosensory acuity in children and adolescents with atypical speech development. Individuals with speech sound disorder (SSD) present with diminished intelligibility in spoken communication that has been associated with negative impacts on social, academic, and socioeconomic outcomes (e.g., Crowe Hall, 1991; Hitchcock et al., 2015; McCormack et al., 2009). When speech deviations persist past the age of 9 years, children may be identified as presenting with residual speech sound disorder (RSSD; Boyce, 2015). Deviations in RSSD tend to cluster on a subset of late-emerging, articulatorily complex sounds such as /ɹ/, /s/, and /tʃ/. RSSD is considered particularly challenging to treat; deviations may persist through adolescence and, in an estimated 2% of speakers, into adulthood (Flipsen, 2015). This study was carried out in conjunction with a broader program of research aimed at improving treatment offerings available to children and adolescents with RSSD affecting /ɹ/ (McAllister et al., 2020), which is sometimes described as the most common residual distortion (Ruscello, 1995).

A sizable body of research has established that sensory feedback plays a crucial role in the development of typical speech production. On average, individuals with RSSD score lower than their peers with typical speech production on measures of both auditory-perceptual acuity (e.g., Cialdella et al., 2021) and oral somatosensory function (e.g., Fucci & Robertson, 1971). However, there is considerable heterogeneity across children and adolescents with RSSD in both domains, including individuals who score in the typical range on perceptual tasks despite the presence of residual distortions in production (Bird & Bishop, 1992; Cialdella et al., 2021).

Measurement of sensory acuity could help guide clinical management of children with RSSD. Here, we take inspiration from a line of second language (L2) learning research sometimes termed the *personalized learning framework* (Perrachione et al., 2011; Wong et al., 2017). This literature posits that L2 acquisition can be optimized by measuring individuals' baseline characteristics and learning outcomes and looking for predictors of learning success. An example of the application of this framework can be found in Perrachione et al. (2011), which engaged native speakers of English in a task of learning to recognize lexical tones. The authors found that success in this task was predicted by performance on a pitch contour perception test (PCPT) administered at baseline. Moreover, they found an interaction between pitch perception ability and training condition: Learners with higher PCPT scores showed better learning outcomes in a training condition characterized by high-variability stimuli as compared to a low-variability condition, whereas learners with poor baseline performance on the PCPT showed significantly lower outcomes in the high-variability compared to the low-variability condition. This study made a strong case that learning outcomes can be maximized when individual differences are taken into consideration in the design of training activities.

This study represents a step toward the application of the personalized learning framework in the context of treatment for RSSD. The long-term goal is to be able to assess children with RSSD, characterize their performance in auditory-perceptual and somatosensory domains, and recommend a treatment approach tailored to their profile of sensory strengths and weaknesses. We discuss this future direction in more detail below. For this clinical application to be possible, however, it is necessary to obtain a detailed characterization of sensory abilities in typically developing children in the same age range. Thus, this study aimed to establish reference data that can accurately and precisely characterize both auditory and oral somatosensory functions in children aged 9–15 years, the age range adopted in numerous previous studies of intervention for RSSD (e.g., Benway et al., 2021; Cialdella et al., 2021; McAllister et al., 2020; Preston et al., 2015). In the sections that follow, we briefly review how auditory-perceptual and somatosensory acuity is measured and review previous literature on maturational changes in both domains.

Auditory-Perceptual Acuity

Measuring Auditory-Perceptual Acuity

Auditory-perceptual acuity for speech refers to an individual's ability to perceive fine differences between speech sounds or to identify speech sound categories given stimuli with varying characteristics. Previous literature suggests that speech perception is subject to considerable variability across typical individuals (e.g., Franken et al., 2017; Kong & Edwards, 2016). Auditory-perceptual acuity has been measured in the literature most frequently by tasks involving identification and/or discrimination of speech sounds (e.g., Gerrits & Schouten, 2004; Perkell, Guenther, et al., 2004; Perkell, Matthies, et al., 2004; Pisoni, 1973). Identification tasks require listeners to label a sound, commonly drawn from a synthetic continuum, as belonging to one of a closed set of phonetic categories. Typically, listeners classify multiple steps along the continuum repeatedly in random order, and their responses are fitted to a logistic function. This fitted function can be used to identify the boundary between phonetic categories (i.e., the point at which either response category is equally likely), as well as the width of the boundary region. Narrower boundary widths are indicative of greater consistency in assigning ambiguous stimuli into phonemic categories (Cheng et al., 2021; Hoonhorst et al., 2011), which has been used as a proxy for auditory-perceptual acuity in some studies (e.g., Cialdella et al., 2021; McAllister Byun & Tiede, 2017). A related task is *category goodness judgment*, in which listeners are presented with diverse stimuli and asked to classify each item as an acceptable or unacceptable instance of a given phonemic category (Dugan et al., 2019; Iverson & Kuhl, 1996; Preston et al., 2015; Tyler, 2021).

Discrimination tasks are designed to measure a listener's ability to identify acoustic differences between sounds. In these tasks, multiple speech stimuli are presented together, and the listener is asked to classify the sounds as the same or different. Again, sounds are commonly drawn from a synthetic continuum, and the distance between the pairs of sounds presented is manipulated to arrive at an estimate of the smallest detectable difference (i.e., just noticeable difference [JND]) for a given listener and contrast. Previous literature has treated performance on discrimination tasks as an index of auditory acuity, with a smaller JND indicating more acute perception (Ghosh et al., 2010; Li et al., 2019).

While the identification and category goodness judgment tasks are inherently phonemic in nature (i.e., they involve classifying speech stimuli relative to categories that can be used to signal a meaningful contrast between words), discrimination tasks can tap either a phonemic mode of perception or an auditory mode in which listeners respond to low-level phonetic detail (Gerrits & Schouten, 2004; Pisoni, 1973, 1975). This study adopted an AXB format, in which listeners hear three speech stimuli in sequence and must determine if the first or last stimulus was different from the second stimulus item (e.g., Li et al., 2019; Perkell, Guenther, et al., 2004; Perkell, Matthies, et al., 2004). Previous literature suggests that performance on this type of task measures both auditory and phonemic levels of perception (Pisoni, 1975).

Developmental Changes in Auditory-Perceptual Acuity

Neural development of the auditory pathway is a protracted process that continues to undergo refinement into late childhood (J. K. Moore & Linthicum, 2007). Numerous behavioral studies have documented changes in auditory-perceptual acuity over the course of maturation (e.g., Dugan et al., 2019; Flege & Eefting, 1986; Hazan & Barrett, 2000). For example, Hazan and Barrett (2000) designed identification tasks for multiple phonemic contrasts (e.g., /s/–/ʃ/) and administered them to children between the ages of 6 and 12 years old (n = 84) and a comparison group of adult speakers (n = 13). They found a statistically significant decrease in boundary width with increasing age. Even the oldest group of children was found to exhibit lower consistency in phonemic classification than the adult comparison group, suggesting that auditory perception continues to undergo refinement well into older childhood. Similarly, Flege and Eefting (1986) investigated changes in boundary location for a plosive voicing contrast in children aged 9–17 years (n = 10) and a comparison group of adults aged 17 years and older (n = 10). They found significant differences in the location of the voice onset time category boundary between 9-year-olds and 17-year-olds, as well as between the 17-year-olds and the adult speaker group, further supporting the notion of a protracted period of refinement of auditory perception that continues into adolescence.

However, other studies have failed to find an association between age and auditory-perceptual acuity. For instance, Hitchcock et al. (2020) found no statistically significant difference between adults (n = 24) and typically developing children aged 7;0–14;0 (years;months; n = 15) on the Wide-Range Acoustic Accuracy Scale, which measures discrimination of three synthetic continua (/bɑ/–/wɑ/, /dɑ/–/gɑ/, and /ɹɑ/–/wɑ/). Other findings suggest that the perceptual task used matters when assessing the relationship between auditory-perceptual acuity and age. Sussman and Carney (1989) found no significant difference between child speakers (n = 30, with 10 participants in each of three age groups) and adults (n = 10) when perception of stop place contrasts was assessed in an identification task. However, when the same contrasts were assessed in a discrimination task, there was a statistically significant difference in performance between the youngest and oldest child groups (5–6 years vs. 9–10 years), as well as between the older children and adults.

Other research suggests that the development of speech perception will also be influenced by characteristics of the phonemic contrast under study, including acoustic salience and the frequency of occurrence of the contrast (Burnham, 1986). Because this study is linked to a program of research investigating RSSD affecting /ɹ/, our auditory-perceptual measures are focused on this phoneme. Previous studies have offered some indication that perception of /ɹ/ continues to mature in late childhood and adolescence. Specifically, in a sample of 40 typically developing children aged 9–14 years, McAllister Byun and Tiede (2017) found a significant negative association between age and boundary width on an identification task with a synthetic continuum from "rake" to "wake." In addition, Dugan et al. (2019) found that responses on a task assessing category goodness judgment for /ɹ/ were more strongly associated with an acoustic measure of rhoticity in an adult listener group (n = 30) than a group of typically developing children aged 7–16 years (n = 35).

Somatosensory Acuity

Measuring Oral Somatosensory Acuity

Somatosensation is commonly described as consisting of both tactile and proprioceptive components (Berryman et al., 2006). Tactile sensation refers to the sense of touch and pressure and arises from the activation of mechanoreceptors when a body structure contacts another surface. Proprioception relates to awareness of the position and movement of body structures and arises from different types of receptors, including muscle spindles and skin mechanoreceptors (Guenther, 2016). In a speech context, proprioception could help speakers identify the direction of movement of speech structures or the relative position of structures, even in the absence of contact between them.

Various methods exist for measuring tactile and/or proprioceptive aspects of somatosensory acuity. Tactile acuity of the speech structures can be assessed through tasks measuring thresholds for detection of touch or vibration (e.g., Etter et al., 2017; Fucci, 1972) or by measuring the smallest distance between points on the tongue surface that the participant perceives as two points rather than one (Engelen et al., 2004; Etter et al., 2017). However, tasks that only require detection of tactile stimuli have been characterized as limiting because they do not assess the function of all types of oral mechanoreceptors (Lee et al., 2022). To overcome this limitation, researchers have recommended the use of tasks measuring oral stereognosis, defined as the ability to identify symbols or forms using the sense of touch with the oral structures (Attanasio, 1987; Fucci & Robertson, 1971). For example, participants may be asked to use the tongue tip to identify embossed letters of different sizes (Lukasewycz & Mennella, 2012; Steele et al., 2014). Lastly, a related task uses blocks engraved with ridges or grating that can differ in spacing; these blocks are presented to the participant, who is asked to identify the orientation of the grating (e.g., horizontal or vertical) using the tongue (Appiani et al., 2020; Lee et al., 2022).

Other literature has investigated the role of proprioceptive acuity in speech production. Some experiments have tapped into proprioceptive acuity by perturbing the position of the articulators with a bite block or palatal prosthesis (e.g., Baum & McFarland, 1997; Gritsyk et al., 2021; Zandipour et al., 2006) or through robotic manipulation of movement trajectories of the articulators (e.g., Feng et al., 2011; Lametti et al., 2012; Nasir & Ostry, 2006; Tremblay et al., 2003). One challenge in measuring the role of proprioception in speech production is the need to eliminate or control for the influence of auditory feedback. Some studies have achieved this through auditory masking (Gritsyk et al., 2021; Zandipour et al., 2006), perturbations that leave speech acoustics unaffected (Nasir & Ostry, 2006; Tremblay et al., 2003), or by measuring formants at an early time point before auditory feedback can be incorporated (Baum & Katz, 1988; Lindblom et al., 1979). Such studies have found that speakers make articulatory adjustments in response to these perturbations even in the absence of auditory feedback, suggesting that somatosensory feedback makes a unique contribution to the planning and control of speech. However, compensation in the presence of masking noise is generally incomplete (Gritsyk et al., 2021; Zandipour et al., 2006).

A distinct approach to the measurement of somatosensory acuity asks participants to explicitly reflect on the position of the articulators while producing different sounds and answer questions such as "Was your tongue in the front or back of your mouth?" (e.g., Gritsyk et al., 2021; Lohman & Fucci, 2000, 2001; Lohman et al., 2001). Lohman et al. (2001) described this as a "lingual-tactile awareness" task and emphasized the contribution of tactile sensation to performance on this task, which is not inappropriate given that their questions focused on tongue placement for lingual consonants (/t, k, ʃ, ɹ, l, θ/). Gritsyk et al. (2021) administered one module focused on consonants and three modules asking participants to characterize the relative position of the articulators when producing pairs of vowels. Given the limited nature of articulator contact during vowel production, Gritsyk et al. characterized their task as primarily assessing proprioceptive aspects of somatosensory function, although a tactile component is undoubtedly also present. This article treats the articulatory awareness task from Gritsyk et al. as a measure of speakers' oral proprioceptive function. This article also draws a contrast between the articulatory awareness task, which measures proprioception in an explicit or metalinguistic fashion, and tasks involving mechanical perturbation of the articulators, where oral proprioceptive feedback is used to adjust articulator position in an implicit or reflexive fashion. To our knowledge, no previous research has compared explicit and implicit measures of oral proprioception in the same speakers.

Developmental Changes in Oral Somatosensory Acuity

Previous literature has also tested for developmental changes in somatosensory function over the course of childhood and early adolescence. In the domain of oral tactile function, several older studies reported an association between age and performance on oral form recognition tasks (Gisel & Schwob, 1988; Kumin et al., 1984; McDonald & Aungst, 1967). However, more recent studies using large sample sizes have failed to find effects of age on tactile acuity. Appiani et al. (2020) examined performance on two oral tactile measures (a grating orientation task and a Von Frey filament task) in a group of children aged 6–13 years (n = 147) and an adult group (n = 70). They found no differences in lingual tactile sensitivity between the adult and child groups or between different age brackets within the child group. Similar results were reported by Lee et al. (2022), who used a grating orientation task to evaluate tactile sensitivity in 1,817 children aged 5–12 years and found no significant differences between age groups (5–7, 8–9, or 10–12 years). It is possible that the differing findings of these two sets of studies reflect the difference in task (oral form recognition vs. grating orientation detection), although the two tasks are thought to assess overlapping aspects of somatosensory function. In addition, Lukasewycz and Mennella (2012) administered an oral letter identification task, which shares many properties with oral form recognition, and found no difference in performance between children aged 7–10 years (n = 52) and adults (n = 46).

Relatively few studies have tested for developmental changes in performance on tasks assessing the proprioceptive aspect of somatosensory function. Previous research investigating maturation of the ability to compensate for physical perturbations such as a bite block has yielded mixed results. Some studies have reported that children compensate for the presence of a bite block as early as 4–5 years of age (Baum & Katz, 1988; DeJarnette, 1988; B. L. Smith & McLean-Muse, 1987), whereas others report a lack of compensation in preschool (Edwards, 1992) or school-age (Gibson & McPhearson, 1980) children. There is a lack of previous research examining compensation for a bite block in the presence of masking noise in child or adolescent speakers.

However, it was argued above that tasks in which participants are asked to reflect explicitly on the position of the articulators in the oral cavity activate proprioceptive and tactile awareness. Lohman et al. (2001) administered a task assessing awareness of lingual placement during consonant production to children in three age groups (mean ages of 6.8, 10.7, and 18.8 years; n = 30 per age group). They found a significant association between age and performance on this task, suggesting that the aspects of somatosensory function assessed by this measure may continue to develop through later childhood and adolescence. However, further research is needed to support the finding of this single study. In addition, the findings reported here suggest that the presence or absence of associations between somatosensory acuity and age may depend on the type of task used to assess somatosensation. This points to a need for further research to assess multiple aspects of somatosensory function and evaluate whether effects of age vary depending on the type of task used.

Possibility of Sensory Trade-Off

A number of previous studies have examined the relationship between acuity in auditory-perceptual and somatosensory domains (Fucci, 1972; Ghosh et al., 2010; Nasir & Ostry, 2008). While it is logically possible that acuity could be positively correlated across domains (i.e., speakers with greater auditory acuity also tend to exhibit greater somatosensory acuity), several studies have found no association between performance on measures of auditory and somatosensory function (Ghosh et al., 2010; D. J. Smith et al., 2020). In fact, some studies have pointed to a negative association between auditory-perceptual and somatosensory domains. In their study evaluating the performance of 100 first-grade children on an oral form discrimination task and an auditory discrimination task, Madison and Fucci (1971) reported a statistically significant negative association, such that children who showed higher auditory acuity performed more poorly in the somatosensory domain. In a study that simultaneously perturbed both auditory and somatosensory feedback during a speaking task, Lametti et al. (2012) found that 21% of participants compensated only for the somatosensory perturbation, 53% only for the auditory perturbation, and 26% for both. Furthermore, they found that participants who compensated more for the auditory perturbation compensated less for the somatosensory compensation, and vice versa. Lametti et al. suggested that their finding could be interpreted as evidence that some speakers display a "sensory preference," weighting either auditory or somatosensory feedback more heavily than feedback in the other domain. Because individual differences in sensory weighting could be relevant for treatment planning, this study followed up on the hypothesis of a possible trade-off between auditory and somatosensory domains in sensory development.

Study Aims

This study collected data from 98 children aged 9–15 years on six measures of perception, three auditory-perceptual and three oral somatosensory. The auditory measures are specific to perception of American English /ɹ/, reflecting our broader program of research investigating RSSD affecting /ɹ/. The somatosensory measures are more general in character because there was no obvious way to link them specifically to /ɹ/. In the context of our wider program of research, the reference data collected here will help to characterize sensory performance in children with RSSD as typical or atypical, with the long-term goal of using this information to personalize treatment recommendations. In the immediate context, this study examined the following questions.

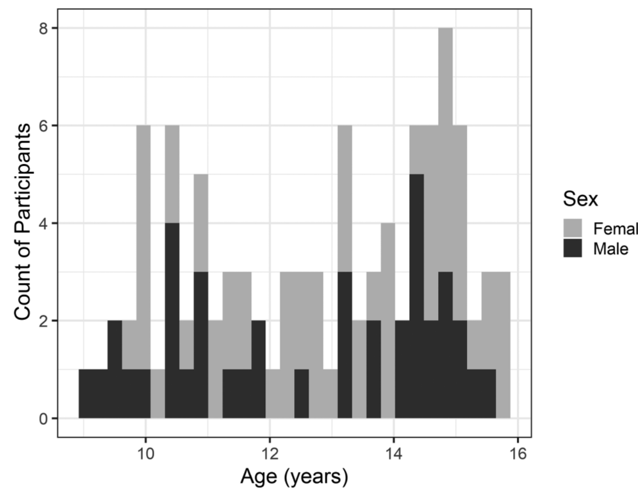
1. Are scores on a given sensory measure significantly associated with scores on the other two tasks designed to test the same domain? This speaks to the concurrent validity of our measures of sensory function. We hypothesized that we would observe at least some statistically significant associations between scores on tasks measuring the same construct. We had less expectation of significant associations between tasks that assess different dimensions of a sensory domain, such as tactile versus proprioceptive aspects of somatosensory function.
2. Are scores on a given sensory measure significantly associated with participant age? Based on previous literature indicating that sensory acuity undergoes gradual refinement into late childhood or even adolescence (Flege & Eefting, 1986; Hazan & Barrett, 2000; Lohman et al., 2001), we hypothesized that at least some measures of acuity in both auditory-perceptual and somatosensory domains would show statistically significant associations with age. We had a greater expectation of associations with age for auditory than somatosensory measures because previous literature provides more robust evidence of late-stage refinement in the former domain; for example, recent research has not found evidence of refinement of oral tactile acuity in older children and adolescents (Appiani et al., 2020; Lee et al., 2022).
3. Do measures of auditory-perceptual and somatosensory acuity support the idea of a "trade-off" between sensory domains? Based on previous literature suggesting that individuals may tend to rely more heavily on one sensory domain or the other (Lametti et al., 2012; Madison & Fucci, 1971), we hypothesized that composite measures of auditory-perceptual and somatosensory acuity would show a statistically significant inverse relationship.

Method

This study was carried out at three sites—New York University, Syracuse University, and Montclair State University. Institutional review board approval was obtained from the Biomedical Research Alliance of New York (Protocol No. 18-10-393). Informed assent was obtained from all participants, as well as written permission from a parent or guardian. All participants were between the ages of 9;0 and 15;11 at the start of the study. This age range was selected based on our goal of providing a point of comparison for children and adolescents with RSSD, since multiple intervention studies targeting this population have adopted the same age cutoffs (e.g., Benway et al., 2021; McAllister et al., 2020). Participants were required to speak English as their dominant or equally dominant language and to have been exposed to English by the age of 3 years, per parent/guardian report. Given the focus of this study on rhotic perception, participants were also required to speak and hear a rhotic variety of English in the home. Participants were excluded if parent report indicated a history of sensorineural hearing loss, major neurological or neurobehavioral disorder (e.g., epilepsy), neurological insult (e.g., brain injury, surgery, or stroke), or developmental disability (e.g., Down syndrome). Participants who met all screening criteria participated in an evaluation session in which they were required to pass a pure-tone hearing screening at 20 dB bilaterally at 500, 1000, 2000, and 4000 Hz. They were also required to achieve a passing score on the Clinical Evaluation of Language Fundamentals–Fifth Edition Screening Test (CELF-5 Screening Test; Wiig et al., 2013) and exhibit grossly typical structure and function in an oral mechanism exam. Finally, participants had to exhibit perceptually accurate production of /ɹ/ sounds on a standard list of 50 words assessing /ɹ/ accuracy and score at or above the 16th percentile for their age on the Goldman-Fristoe Test of Articulation–Third Edition (GFTA-3; Goldman & Fristoe, 2015). A total of 119 participants were initially recruited to the study. After the application of the criteria listed here, a total of 98 children were included. Fifty-eight of the participants were female (59.2%), indicating a slight bias toward sampling from female participants. Table 1 shows descriptive characteristics for included participants (mean, standard deviation, and range for GFTA-3 standard score and CELF-5 Screening Test raw score); individual-level characteristics are reported in a table in the supplementary materials (https://osf.io/8kvp4/). In addition, Figure 1 represents the distribution of participants across the range of age in years, shaded by sex. Figure 1 shows that the younger end of the age range is undersampled relative to the older end. The less-than-ideal distribution of participants across age and sex categories can be partly attributed to the challenges of recruiting during the COVID-19 pandemic: In particular, the overrepresentation of older children reflects the fact that children under 12 years were not eligible for vaccination and therefore less likely to participate in in-person research studies for a significant fraction of the study period.

We administered a battery of auditory and somatosensory measures described in detail below. Auditory-perceptual tasks included (a) identification of items along a synthetic continuum from “rake” to “wake,” (b) discrimination of stimulus pairs from the same continuum in an AXB task, and (c) category goodness judgment for words containing /ɹ/ produced by various speakers with and without RSSD affecting /ɹ/. Somatosensory tasks included (a) an oral stereognosis task measuring tactile input received by the articulators (Steele et al., 2014), (b) an articulatory awareness task probing the sense of tongue position (Gritsyk et al., 2021; Lohman et al., 2001), and (c) a bite block task measuring the ability to compensate for an articulatory perturbation in the presence of masking noise (Gritsyk et al., 2021; Zandipour et al., 2006). Stimulus materials, experimental scripts, and protocols to reproduce these tasks are available on the Open Science Framework at https://osf.io/bkasm/.

Figure 1. Histogram of participants across the range of age in years, shaded based on sex.



Due to COVID-19 masking requirements, which varied over time and across the three sites involved in the study, not all participants were able to complete certain somatosensory tasks that required participants to be unmasked. The bite block task was completed by 95 out of 98 participants, and the oral stereognosis task was completed by 77 out of 98 participants.

Because this study involved administration of audio stimuli using different equipment across three research sites, an initial calibration phase was undertaken to identify settings that would yield equivalent output loudness across devices. All sites used a Sound BlasterX G5 external

Table 1. Characteristics of included participants.

Measure	*M*	*SD*	Minimum	Maximum
Age (years)	12.8	2.0	9.0	15.8
GFTA-3 (standard score)	103.9	2.1	98.0	110.0
CELF-5 (raw score)	24.2	3.5	14.0	31.0

Note. GFTA-3 = Goldman-Fristoe Test of Articulation–Third Edition; CELF-5 = Clinical Evaluation of Language Fundamentals–Fifth Edition Screening Test.

sound card when presenting audio stimuli. For calibration, the insert earphones used for audio presentation in the bite block task were coupled to a sound pressure level meter via an occluded ear simulator. The device settings for each laptop were adjusted until the achieved intensity of audio presentation via the Sound BlasterX G5 card matched the target intensity of 78 dB, as well as step-down frequencies of 75 and 72 dB (for participants who could not tolerate presentation at 78 dB). These device settings were noted in site-specific protocols and used throughout the administration of the bite block task. A standard playback level of 78 dB was used for presentation of stimuli for the identification, discrimination, and category goodness judgment tasks. Participants completed all tasks in a sound-shielded booth or quiet room. Stimuli for all three auditory-perceptual tasks were presented via HD 280 PRO Sennheiser headphones.

Auditory-Perceptual Measures

Identification

A computerized task, adapted from McAllister Byun and Tiede (2017), was used to assess the consistency with which listeners partitioned a synthetic stimulus continuum into /ɹ/ and /w/ phoneme categories. Productions of the words "rake" and "wake" were elicited from typically developing children in the same age range as our target participants. STRAIGHT synthesis was used to generate 240-step continua between end-point tokens that were matched with respect to amplitude and timing of the voiced segments and burst release onsets. All tokens were normalized to a root-mean-square amplitude of 70 dB. Members of the study team reviewed these preliminary continua for considerations such as clarity of intermediate tokens and balance of the continuum around the perceived phoneme boundary. Based on these impressions, a single continuum generated from productions of a 13-year-old female talker was selected for use in the study. The 240-step continuum was then reviewed with the goal of selecting a subset of steps that spanned the range between the end points and were roughly equally spaced with respect to the third formant (F3), which makes a crucial contribution to the perceptual difference between /ɹ/ and /w/. F3 frequencies were both Bark- and log-transformed to achieve more linear spacing between continuum steps. The transformed data were plotted with continuum step on the *x*-axis and transformed F3 frequency on the *y*-axis with a superimposed best-fit line; the experimenters manually selected 12 tokens that were close to the line and roughly equally spaced. The first and 12th tokens were set aside for training purposes, and continuum Steps 2–11 formed the basis for the 10-step continuum that was presented to participants in the main experimental task.

Stimulus presentation and response recording were computerized using custom software. In each trial, participants heard a single token from the continuum presented twice in a row with a 200-ms interval separating the two repetitions; the goal of this double presentation was to increase the robustness of the perceptual task to fluctuations in participant attention. Participants were instructed to identify each token pair as "rake" or "wake" in a forced-choice task by clicking on the corresponding text on the computer screen. In an initial practice phase, participants completed five trials featuring the two end-point tokens (Steps 1 and 12). If they scored 5/5 correct, they advanced to the main trials; if not, the instructions and the practice trials were repeated up to 2 more times, whereupon all participants were advanced to the main trials. In the main task, participants heard Steps 2–11 eight times each, for a total of 80 trials. Stimuli were randomly ordered in a cyclic fashion, such that all 10 continuum steps were presented in random order before any step was repeated. A brief break was offered after every 20 trials.

To analyze the results for each participant, the proportion of tokens identified as "rake" was plotted for each of the 10 steps of the main continuum. (Practice trials featuring Steps 1 and 12 were not included in this analysis.) These data points were fitted to a logistic function via maximum likelihood estimation. Auditory-perceptual acuity for /ɹ/ was calculated as the width of the fitted function from the 25th to the 75th percentile of probability. A narrower boundary region is indicative of higher response consistency, and this was interpreted as evidence of more acute perception of the contrast in question (Hazan & Barrett, 2000; McAllister Byun & Tiede, 2017).

Discrimination

An auditory discrimination task implemented in MATLAB was developed using the same "rake"–"wake" stimulus continuum as the identification task. The location of the boundary between /ɹ/ and /w/ was set in an individualized fashion based on the location of the 50th percentile of probability in the logistic function fitted to participants' stimulus identification responses. This value (expressed in continuum steps) was entered into the MATLAB program, which then used linear interpolation to find the nearest stimulus in the full 240-step continuum. This stimulus was then used as a central reference point in the AXB discrimination task. In the task, participants heard three tokens from the stimulus continuum in sequence, separated by 500 ms. The first and third tokens consisted of a token that was a fixed distance along the continuum below the central reference point and a token that was the same distance above the central reference point, with the order randomized. The second (central) token always corresponded to the lower of the two stimuli. Hence, on any given trial, either the first or third token was different from the central token. To maintain the engagement of young participants in a potentially lengthy and demanding task, stimuli were presented in a gamified interface. Three characters were shown on the screen who “produced” the three tokens in sequence; participants were challenged to “find the impostor” whose production did not match the character in the middle. Participants received an animated visual reinforcement and were awarded a point in the game for each correct response.

For all participants, the experiment began with an initial separation of 40 continuum steps that was identified through pilot testing. In subsequent trials, a one-up, two-down procedure was used to estimate JND for each participant in an adaptive fashion. At the beginning of the procedure, stimulus distance was decreased by eight steps following a correct response and increased by four steps following an incorrect response. After every fourth reversal (i.e., change from correct to incorrect responses, or vice versa), the size of the adjustment in distance between stimuli was reduced by half. The task was discontinued after either 12 reversals or 80 trials, whichever came first. JND was calculated by obtaining the mean of the distance between the stimuli at the final four reversals. A smaller JND is interpreted as consistent with a higher degree of auditory-perceptual acuity.

Category Goodness Judgment

Finally, participants’ auditory-perceptual acuity was assessed in a task of classifying various speakers’ productions of /ɹ/ as correct or incorrect. Stimuli were 100 naturally produced single words containing /ɹ/ in various word positions (no more than one /ɹ/ per word). These utterances were elicited from a total of 52 speakers including both children and adults with both typical speech and RSSD affecting /ɹ/. All were speakers of a rhotic dialect of English. Each /ɹ/ token was identified as “correct” or “incorrect” based on consensus across at least four expert listeners; tokens where consensus could not be reached were excluded. Fifty items had an intended response of “correct,” and 50 had an intended response of “incorrect.” Twenty-eight items featured syllabic /ɹ/, 22 featured post-vocalic /ɹ/, 25 featured singleton /ɹ/ in onset position, and 25 featured /ɹ/ in onset position in a consonant cluster.

Participants were informed that they would hear words containing the /ɹ/ sound and that their task was to “choose whether the /ɹ/ sound is right or wrong” by clicking a button with the corresponding label. Stimulus presentation and response recording were carried out using Praat software (Boersma & Weenink, 2019). No training was provided in this relatively self-explanatory task. A brief break was offered after every 25 trials. Performance was assessed by computing the number of correct responses (i.e., where the participant’s response agreed with the consensus judgment across trained listeners) divided by the total number of trials.

Oral Somatosensory Measures

Oral Stereognosis

This task measures oral tactile acuity in a nonspeech task by asking participants to identify raised letter forms embossed on Teflon strips using their tongue. The participants held the letter strips and could actively search each letter shape with the tongue tip. The plastic strips were presented face down, and participants wore sunglasses to further ensure that they did not see the letter prior to insertion in the oral cavity. Participants were provided with diagrams and verbal explanation illustrating the orientation of the letters (i.e., top of letter toward back of mouth) and were required to demonstrate understanding of this information prior to proceeding to the main task. The protocol and stimulus materials used in this study followed the specifications in Steele et al. (2014). The letters ranged in size from 2.5 to 8 mm. All letters were capitalized in a sans serif font and were drawn from the set {A, I, J, L, T, U, W}. The task followed an adaptive staircase design in which incorrect responses resulted in a presentation of a larger letter size and correct responses resulted in a smaller letter size. The task ended after eight reversals (changes from correct to incorrect or incorrect to correct) or 28 trials, whichever came first. All responses were recorded in a score form by the examiner. Following Steele et al., the final score was calculated as the average letter size across all correct responses.

Articulatory Awareness

The articulatory awareness task, developed in Gritsyk et al. (2021), was designed to tap into participants’ sense of proprioception within the oral cavity. Throughout the task, participants were prompted to produce a variety of speech sounds and answer questions about the relative position of their articulators. PsychoPy 3 (Version 3.2.3; Peirce et al., 2019) was used to administer the task. The participants listened to prerecorded prompts via HD 280 PRO Sennheiser headphones while viewing corresponding text on the screen of the laptop used to administer the task. The task consisted of four parts with eight to nine stimuli per block. In each block, target sounds were modeled verbally and orthographically in isolation as well as within a real word (e.g., “ee” like in “heat”). In order to help the participant gauge the sensation of tongue position during production of the sounds of interest, participants were prompted to repeat each target sound in isolation 3 times. In the first block, participants were asked to indicate if they felt that a given consonant sound (e.g., /t/) was produced with the front or the back of the tongue. For the remaining three blocks, participants were prompted to produce two vowel sounds in an alternating fashion 3 times (e.g., /u/–/i/, /u/–/i/, /u/–/i/). They were then asked to report the relative position of the tongue during production of the sounds in a pair (e.g., "Which sound was further back/higher/lower?"). Responses were scored based on the relative position of vowels in F1–F2 space across regional dialects of American English (Clopper et al., 2005); pairs of vowels that occupy overlapping regions of acoustic space were avoided. In the event that a vowel was produced incorrectly, one additional verbal model and prompt to repeat was provided. If the sound continued to be produced in error, the task proceeded, and the participant's actual production was phonetically transcribed for reference. All responses were provided verbally by the participant and entered into the experiment by the examiner. For this task, the primary outcome measure was the percentage of correct responses across all sections of the task (34 items in total).

Bite Block With Auditory Masking

In bite block tasks, a physical perturbation is introduced into the oral cavity, and the experimenter measures the extent to which the participant is able to approximate typical production of a speech target in the presence of this obstacle. In a bite block task with auditory masking, the participant must attempt to compensate for the perturbation with a minimum of auditory feedback. Thus, the task aims to measure the participant's ability to tune in to somatosensory feedback to determine the actual position of the articulators and adjust their articulatory movements in an effort to offset the presence of the perturbation. In this study, the bite block task was administered in multiple phases. In the first phase (baseline), participants were familiarized with the written words "heed" /hid/, "who'd" /hud/, "had" /hæd/, and "hod" /hɑd/ and then were prompted to repeat each word 8 times each for a total of 32 productions. During the baseline phase, participants heard auditory masking noise (described in detail below), but no articulatory perturbation was applied. The target words were presented on a PC monitor in a pseudorandomized order using custom software. During the following two phases, participants read the same words with masking noise while holding a tongue depressor with the front incisors in one of two orientations. In the "horizontal" bite block condition, participants were cued to produce words containing the low vowels /æ/ and /ɑ/ ("had" and "hod") with the tongue depressor placed horizontally between the front incisors, prompting a nearly closed jaw position. In the "vertical" bite block condition, participants produced words containing the high vowels /i/ and /u/ ("heed" and "who'd") with the tongue depressor placed vertically between the front incisors, creating a jaw aperture of 1.75 cm (the width of the tongue depressor). Each phase of production with a bite block (horizontal and vertical) elicited 16 repetitions of each word in randomized order, for a total of 32 words in each condition. The order of elicitation of the horizontal versus vertical condition was randomly determined for each participant.

To minimize participants' access to auditory feedback during the task, two types of masking noise were presented through two channels. Pink noise was presented through Bluetooth-enabled bone-conduction headphones (Z8) via Audacity (Audacity Team, 2019), and multitalker babble was presented via Praat (Boersma & Weenink, 2019) through in-the-ear noise-isolating insert headphones (Etymotic Research HF5). The pink noise was played at the maximum device volume in the Z8 bone-conduction headphones, and the multitalker babble was played at 78 dB in the insert headphones.[1] As an additional measure to limit auditory feedback, participants were cued to maintain their volume at a low level (quietest possible phonation without whispering) using visual feedback from a digital sound level meter (Kay Pentax, Computerized Speech Lab, Model 4500). The experimenter also monitored the intensity of productions and prompted participants to repeat any tokens judged to be too loud (and therefore likely to provide auditory feedback) or too soft to yield measurable formants. Productions were recorded to a Marantz digital recorder through a tabletop microphone (Shure SM48) positioned 5 in. from the participant's mouth. A sampling rate of 44100 Hz and 16-bit encoding were used.

Due to a miscommunication across study sites, one stimulus ("hod," intended pronunciation /hɑd/) was initially elicited with an /oʊ/ vowel at one site and an /ɑ/ vowel at the others. Once this inconsistency was detected, all five vowels were elicited from subsequent participants.[2] The data from these participants who produced all five vowels (n = 37) were used to impute /ɑ/ values for the participants who lacked that vowel (n = 18). Specifically, a linear regression was fit with measures of the vowels /i/, /u/, /æ/, and /oʊ/ as predictors, as well as the interaction between /æ/ and /oʊ/ (the two vowels in the horizontal

[1]Because some children are sensitive to loud noise, participants were asked if the 78 dB volume was uncomfortably loud, and the playback volume was lowered to 75 dB if they responded affirmatively. A total of six participants completed the task with a playback volume of 75 dB; 78 dB playback was used for the remaining 89 participants.

[2]To elicit the extra vowel without altering the procedures or number of trials for the other vowels, a second baseline and additional bite block condition eliciting /oʊ/ were added after completion of the original conditions described above. The second baseline consisted of the original four target words plus the additional word "hoed" /hoʊd/. The second bite block condition (horizontal only) elicited 16 productions of the word "hoed," intermixed with a total of 16 additional productions of "had" and "hod"; these were treated as foils and not analyzed.

condition besides /ɑ/), plus participant age and sex. To improve model fit, a square root transformation of the outcome was used. The overall model was significant, $F(7, 36) = 4.2$, $p = .002$, with $R^2 = .45$ and adjusted $R^2 = .35$. The fitted model was then used to predict /ɑ/ values for the participants who lacked /ɑ/ but possessed all the other measured values (/i/, /u/, /æ/, /oʊ/, age, and sex). To reflect the sampling variability expected in real data, a deviation from the prediction for each imputed value was added, drawn from a normal distribution with a mean of 0 and a standard deviation equal to the residual standard deviation from the regression model. The predictions were then transformed back to the original scale of the data. Because only the corner vowels /i/, /u/, /æ/, and /ɑ/ were used in the final measure of performance on the bite block task, no imputation was undertaken for the participants (n = 30) who possessed measures for the corner vowels but not /oʊ/.

The results of the bite block task were analyzed as described in the study of Gritsyk et al. (2021). The first and second formants (F1 and F2) of all vowels in the baseline and bite block conditions were measured using the Burg method in Praat. Members of the research team examined each participant's sound file in Praat and selected the most appropriate formant settings (five formants in 4500, 5000, 5500, or 6000 Hz) by comparing the automated formant tracking with the visible energy bands of the spectrogram. A single formant setting was selected to use across /u/, /ɑ/, /æ/, and /oʊ/ (when relevant) for each participant. The process was repeated separately to select the best formant settings for /i/, which was noted to require different settings (five formants in 6000, 6500, 7000, or 7500 Hz) due to the high frequency of F2. Using the selected settings, the researchers identified a point within the steady state of the vowel,[3] avoiding regions of poor formant tracking. Formant values were extracted from these steady states using a 50-ms Gaussian window and Bark-transformed using the "vowels" package (Kendall et al., 2018) in the R software environment (R Core Team, 2019). Next, mean F1 and F2 values were calculated for each vowel in the baseline condition. We measured the effect of the bite block on vowel production by calculating the Euclidean distance (ED) in F1–F2 space between each vowel production in the bite block condition and the mean values for the same vowel in the baseline condition. We then obtained the mean ED across tokens for each vowel in the bite block condition for each participant. The mean ED across the vowels /i/, /u/, /a/, and /æ/ was used as the summary measure of bite block performance for each participant, where a smaller ED is suggestive of greater compensation for the bite block perturbation.

To check for errors in formant tracking, plots of ED in baseline and bite block conditions were generated for each vowel and participant. Two of the authors visually inspected these plots to identify points that may have been affected by errors in formant tracking (i.e., tokens that visually appeared as outliers with respect to ED relative to surrounding productions). These tokens were flagged, and the spectrogram was revisited to identify if the unusual ED value might be attributable to poor formant tracking. If formant tracking in the selected window was judged to be of poor quality, the author selected a more representative window in the production. If it was not possible to identify a steady-state window with acceptable formant tracking, the token was discarded (n = 8).

To assess reliability, the fourth author followed the same procedure for 15 of 85 participants (17.6%), measuring one condition (baseline, vertical bite block, or horizontal bite block) per participant. Intraclass correlation (ICC) with single random raters was used to assess the reliability of F1 and F2 measurements between the original and remeasured files. For F1, ICC was computed to be .94 (95% confidence interval [.93, .94]). For F2, ICC was computed to be .95 (95% confidence interval [.94, .96]). These values indicate excellent reliability between formant measurements carried out by different individuals (Koo & Li, 2016).

Analyses

Study data were stored using REDCap secure data capture tools hosted at Syracuse University (Harris et al., 2009, 2019), with all values double-entered to minimize the likelihood of errors. Members of the study team reviewed the records of participant performance on all sensory measures and flagged any items that were suggestive of error (e.g., random response patterns suggestive of inattention) for further review and possible exclusion. Distributions of observations were visually inspected, and it was determined that not all measures yielded normally distributed data. Therefore, nonparametric statistics were used for measures of central tendency and association between variables. Observations were identified as outliers if they fell at least three median absolute deviations (MADs) from the group median (Leys et al., 2013). After the removal of outliers (see Results section for the number of exclusions), Spearman's rho were used to examine associations between variables. This included pairwise associations between the three measures in each domain, as well as associations between each measure and age. Within each domain (auditory and somatosensory), p values were adjusted for multiple comparisons using the false

[3]For the diphthong /oʊ/, the most stable portion of the onglide was used.

discovery rate correction. For the final research question, scores on each task were expressed in terms of positive or negative MAD from an age-specific median, and this transformed measure was averaged across tasks within each domain. The association between average scores in auditory and somatosensory domains was examined using Spearman's rho. Complete code and deidentified data to reproduce the results reported here can be retrieved on the Open Science Framework at https://osf.io/2hy65/.

Results

Auditory-Perceptual Measures

A total of three observations (two identification and one discrimination) were judged to reflect performance suggestive of inattention and were excluded prior to analysis. An additional three observations were dropped as outliers (≥ 3 MADs from the group median), all based on JND on the discrimination task.

For the identification task, the median location of the boundary fell at 6.0 (MAD = 0.4) along the 10-step continuum. The median boundary width was 0.7 (MAD = 1.1). As in previous studies utilizing this task (Cialdella et al., 2021; McAllister Byun & Tiede, 2017), a ceiling effect was apparent; a boundary width of 0.0 was calculated for 39 out of 98 participants. For the discrimination task, the median JND was 10.5 steps based on a 240-step continuum (MAD = 5.9).

For the category goodness judgment task, the median overall percent correct was 84.0 (MAD = 10.4). We examined the breakdown of accuracy across various subparts of the category goodness judgment task. Descriptively, it was noted that participants had higher agreement with experienced listeners when affirming correct tokens as correct (median 98.0% correct, MAD = 3.0) as opposed to rejecting tokens as incorrect (median 72.0% correct, MAD = 23.7). Performance differed minimally across consonantal versus vocalic (i.e., syllabic or postvocalic) variants of /ɹ/ (for consonantal, median = 83.3% correct, MAD = 11.0; for vocalic, median = 86.5% correct, MAD = 11.4). However, accuracy was lower for /ɹ/ sounds that appeared as part of a consonant cluster (median = 77.8% correct, MAD = 11.0) than /ɹ/ sounds in noncluster contexts (median = 87.3% correct, MAD = 9.4).

Pairwise associations were examined between the auditory-perceptual measures (identification boundary width, discrimination JND, and category goodness judgment percent correct), as well as between each task and participant age. The association between identification boundary width and discrimination JND, Spearman's rho(91) = 0.25, p = .04, was statistically significant after correction for multiple comparisons. Consistent with our hypothesis, listeners with wider identification boundaries (suggestive of less acute perception) also tended to have a larger discrimination JND. There was also a statistically significant association between age and percent correct on the category goodness judgment task, Spearman's rho(93) = 0.32, p = .01, indicating that older participants tended to be more accurate than younger participants in classifying /ɹ/ sounds produced by various speakers. Both associations are weak in magnitude, with substantial heterogeneity across individuals (see Figure 2); some children under 10 years achieve scores on par with the oldest age group, and vice versa.[4] No other associations were statistically significantly different from zero. A table reporting all associations can be found in the supplementary materials (https://osf.io/8kvp4/).

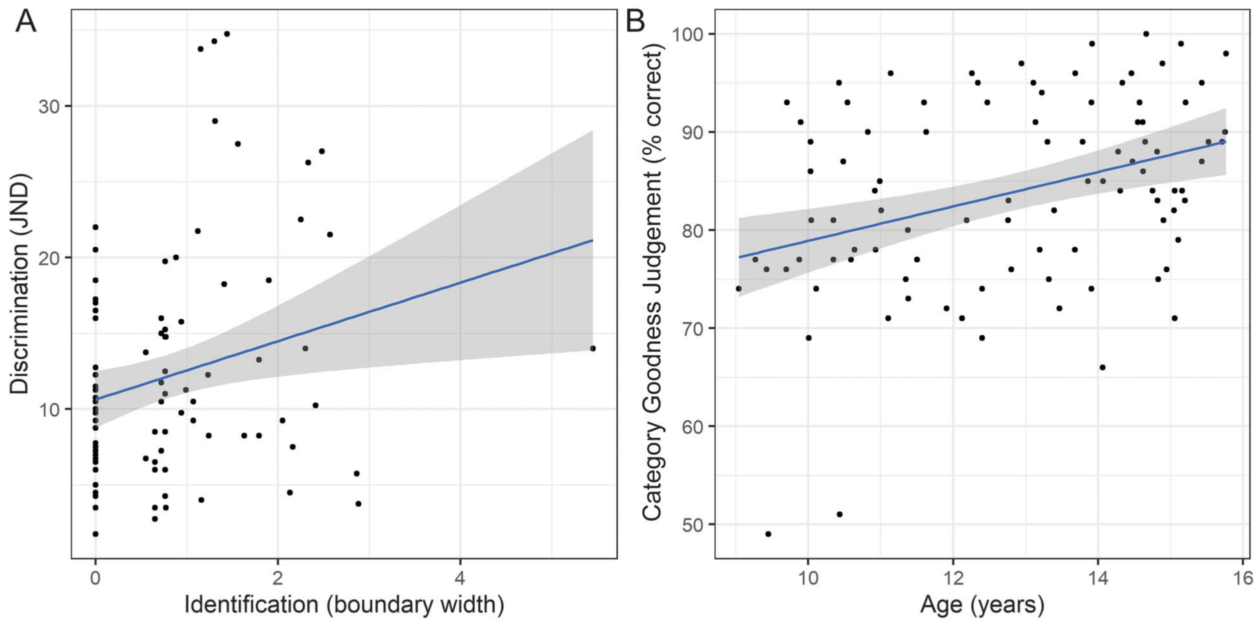
Oral Somatosensory Measures

A total of two observations (both from the oral stereognosis task) were affected by experimenter error and excluded prior to analysis. In addition, three scores were dropped as outliers (> 3 MADs from the group median) with respect to performance on the bite block task.

For the stereognosis task, the median letter size across all correct trials was 5.5 mm (MAD = 1.6 mm). For the articulatory awareness task, the overall median percent correct was 73.5 (MAD = 13.1). For the submodules of the articulatory awareness task, participants were most accurate on the "front or back consonants" module (median percent correct = 77.8, MAD = 16.5) and least accurate on the "which vowel is further back" module (median percent correct = 55.6, MAD = 32.9). The "which vowel is higher" and "which vowel is lower" modules asked the same question in different ways as a means to assess the reliability of the measure. These two modules were associated with very similar levels of accuracy (median percent correct = 75.0, MAD = 37.1 for the "which vowel is higher" module; median percent correct = 75.0, MAD = 18.5 for the "which vowel is lower" module). Scores on the two modules were statistically significantly associated with a moderate magnitude, Spearman's rho(96) = 0.43, p < .001. Finally, for the bite block task, the median ED from baseline formant frequencies was 0.93 (MAD = 0.38) when averaging across the vowels /i/, /u/, /æ/, and /ɑ/ in both perturbed conditions. In the

[4]In the data for the identification task, one participant presents visually as an outlier, although they were not flagged as such on the basis of median absolute deviation from the group median. With the exclusion of this observation, the association between identification boundary width and discrimination JND remains similar in magnitude but ceases to reach statistical significance after correction for multiple comparisons, Spearman's rho(90) = 0.24, p = .05.

Figure 2. (A) Association between boundary width on the identification task and just noticeable difference (JND) on the discrimination task. (B) Association between percent correct on the category goodness judgment task and age in years. Shaded band represents a 95% confidence interval around the best-fit line.



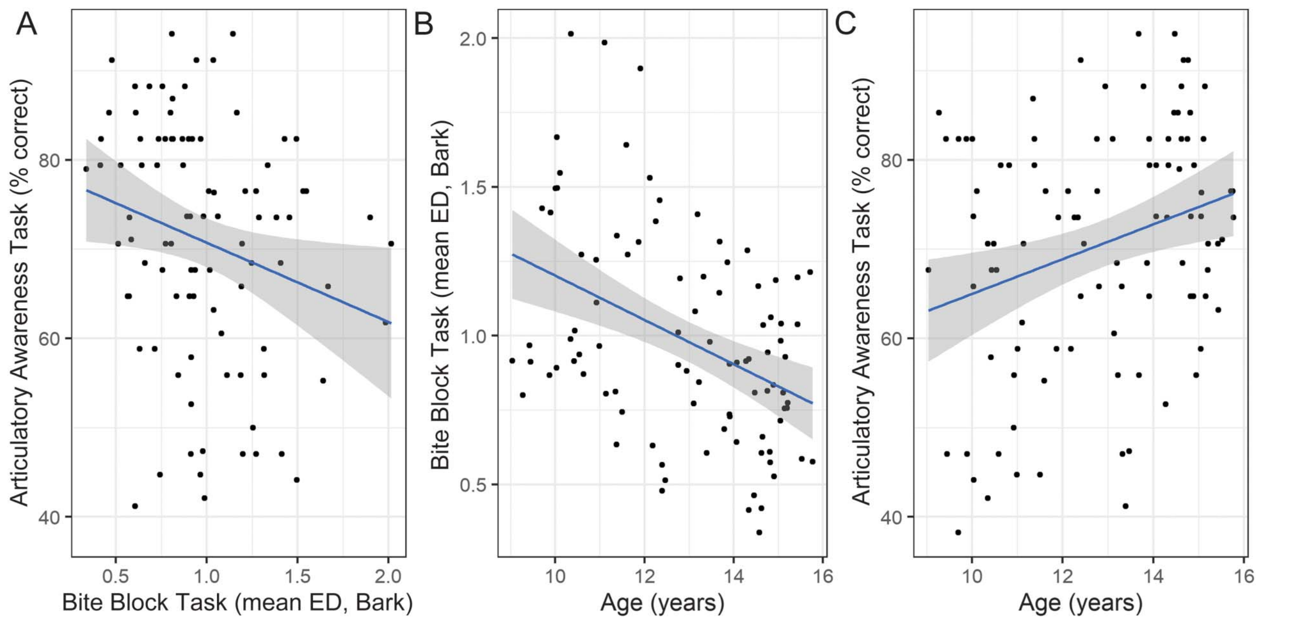
horizontal condition (/æ/, /ɑ/), the median ED was 0.85 (MAD = 0.42), while in the vertical condition (/i/, /u/), the median ED was 0.97 (MAD = 0.41).

Pairwise associations were examined between the oral somatosensory measures (median letter size on the oral stereognosis task, percent correct on the articulatory awareness task, and overall mean ED in Bark between baseline and perturbed tokens for the bite block task), as well as between each task and participant age. The association between scores on the bite block task and the articulatory awareness task was statistically significant after correction for multiple comparisons, Spearman's rho(90) = −0.28, p = .02. The negative direction of the association indicates that smaller ED scores, which are suggestive of better compensation for the bite block, correlated with higher articulatory awareness. There were no statistically significant associations between the stereognosis task and the other tasks. Age was statistically significantly associated with overall performance on the bite block task, Spearman's rho(90) = −0.38, p = .001. The association between age and performance on the articulatory awareness task, Spearman's rho(96) = 0.22, p = .06, did not survive correction for multiple comparisons. In both cases, the direction of the association indicates that older participants tended to show higher somatosensory acuity than younger participants. However, both associations are weak in magnitude, as evidenced by the wide variation across participants shown in Figure 3. Finally, there was insufficient evidence to reject the null hypothesis of no association between age and performance on the stereognosis task at a 5% significance level. However, stereognosis data were available for only 75 participants, giving this test lower power than the other associations examined. A table reporting all associations is provided in the supplementary materials (https://osf.io/8kvp4/).

Possible Trade-Off

A final analysis tested the possibility of a trade-off between auditory-perceptual and somatosensory acuity. For comparison across tasks, absolute scores were converted into relative measures. Because scores on multiple tasks were found to covary with age, relative scores were computed in an age-specific fashion. Model-based clustering was used to select empirically motivated age groupings. Age and scores on all tasks that showed a statistically significant association with age (i.e., the category goodness judgment, articulatory awareness, and bite block tasks) were included in an unsupervised model-based clustering analysis using the "mclust" package (Scrucca et al., 2016). The optimal solution found two clusters with a cut point at 12.5 years of age. Therefore, children were partitioned into a "younger" group (9–12.5 years, n = 43) and an "older" group (above 12.5 years, n = 55). Scores were then transformed to reflect positive or negative deviation from the measure of central tendency for each child's age group. Due to the asymmetric nature of our data, these

Figure 3. (A) Association between Euclidean distance (ED) between bite block and baseline conditions from the bite block task and percent correct on the articulatory awareness task. (B) Association between age in years and ED from the bite block task. (C) Association between age in years and percent correct on the articulatory awareness task. Shaded band represents a 95% confidence interval around the best-fit line.



relative scores were expressed in terms of the number of MADs above or below the group median. Median and MAD scores for each task and age group are reported in Table 2.

The measures in our data set also vary in directionality (i.e., whether a higher or lower score represents stronger performance). Thus, scores on the identification, discrimination, stereognosis, and bite block tasks were reverse-coded so that for all scores, a larger value represents stronger performance. Then, the average was computed across the tasks making up each domain, yielding a single mean auditory-perceptual score and a single mean somatosensory score for each participant.[5] There was no significant association between mean auditory-perceptual acuity and mean somatosensory acuity across participants, Spearman's rho(96) = 0.12, p = .23.

Discussion

The analyses reported here revealed a small but statistically significant association between two measures in the auditory-perceptual domain (identification and discrimination), as well as between two measures in the somatosensory domain (bite block with auditory masking and articulatory awareness). These findings support the concurrent validity of the measures adopted here, although the small magnitude of the observed associations limits the strength of the conclusions that can be drawn. Statistically significant associations with age were found for one auditory-perceptual measure (the category goodness judgment task) and one somatosensory measure (the bite block task), as well as a trend toward an association with age for the articulatory awareness task that did not survive correction for multiple comparisons. There was no evidence of a trade-off between acuity in auditory-perceptual and somatosensory domains. Here, we explore possible explanations and implications of these findings, as well as directions for future research. We also acknowledge factors other than raw sensory acuity that may have influenced these findings, most notably the capacity to sustain attention to experimental tasks.

Auditory-Perceptual Measures

Boundary width on the identification task and JND on the discrimination task were significantly associated. While previous research suggests that identification tasks primarily assess the phonemic level of perception and discrimination tasks assess the auditory level, it has been suggested that the ABX design adopted in the present discrimination task measures both auditory and phonemic

[5]Missing values were dropped in computing this average. Thus, if a participant completed the articulatory awareness and bite block tasks but not the stereognosis task, the average score would reflect the mean across the two tasks that were completed.

Table 2. Median and median absolute deviation scores on each task, partitioned by age.

Age group	Identification boundary width (continuum steps)[a]	Discrimination JND (continuum steps)[b]	Category goodness judgment (percent correct)	Stereognosis (mean letter size in mm)	Phonetic awareness (percent correct)	Bite block (Euclidean distance from baseline in Bark)
Younger (≤ 12.5 years)	0.72 (1.07)	10.25 (5.56)	79 (9.64)	5.69 (1.5)	70.59 (17.44)	1.02 (0.44)
Older (> 12.5 years)	0.65 (0.96)	10.62 (5.56)	87 (8.9)	5.13 (1.47)	73.68 (12.85)	0.88 (0.27)

Note. JND = just noticeable difference.
[a]Relative to a 10-step continuum. [b]Relative to a 240-step continuum.

levels of perception (Pisoni, 1975). The fact that stimuli for both tasks were drawn from the same synthetic continuum from "rake" to "wake" likely supported the finding of significant association between the tasks. The category goodness judgment task also assessed children's perception of American English /ɹ/, but it differed from the identification and discrimination tasks in various ways. It involved naturally produced rather than synthetically modified speech tokens and presented /ɹ/ in multiple phonetic contexts (e.g., in consonant clusters or in postvocalic position) rather than in a single prevocalic context. It also featured diverse talkers whose productions of /ɹ/ could take various forms, including sound substitutions (e.g., [w] for /ɹ/) as well as distortions. These differences set the category goodness measure apart from the other two tasks and could offer a partial explanation for the lack of a significant association. Still, because all the tasks were designed to assess the robustness of participants' representation of the phonetic properties of /ɹ/, the absence of any association is surprising.

One possibility is that our choice of summary measure for the category goodness task was not ideal. The use of overall percent correct could mask differences between children who brought different strategies to the task of identifying correct and incorrect productions. In particular, children in our study tended to have more difficulty rejecting distorted tokens versus endorsing typical productions of /ɹ/ as perceptually accurate (see also Dugan et al., 2019). One way to address this issue is to use signal detection measures such as d' or its nonparametric counterpart A'. As a follow-up analysis, we converted our percent correct scores to the nonparametric A' measure. However, A' was very closely associated with percent correct (Spearman's rho(91) = 0.99, $p < .001$), and substituting A' did not change the significance or direction of any of the associations reported on the basis of percent correct. The percent correct measures also were reasonably well dispersed over a range from 70% to 100% and were roughly normally distributed, which suggests that we cannot point to limited variance or ceiling effects on the category goodness measure as a possible explanation for the lack of a significant association. On the other hand, the identification task did show a clear ceiling effect and a nonnormal distribution; we return to this point under Limitations below.

In our second research question, we hypothesized that we would observe an association between age and performance on at least some of our auditory-perceptual acuity measures. Several previous studies have reported that performance on a variety of speech perception tasks improves with age, with ongoing refinement into adolescence or even early adulthood (Flege & Eefting, 1986; Hazan & Barrett, 2000; McAllister Byun & Tiede, 2017). On the other hand, previous literature has documented extensive individual variability in auditory processing throughout the course of late childhood (D. R. Moore et al., 2011). In this study, a statistically significant association was observed between age and performance on the category goodness judgment task. This aligns with the results of the previous study of category goodness judgment for /ɹ/ by Dugan et al. (2019). Both studies support the hypothesis that the ability to classify diverse speakers' productions of /ɹ/ as perceptually "correct" or "incorrect" is subject to ongoing refinement in later childhood.

We did not observe a statistically significant association between age and performance on tasks of identifying and discriminating stimuli from the synthetic "rake"–"wake" continuum. The lack of a significant association with performance on the identification task was unexpected because McAllister Byun and Tiede (2017) found a small but statistically significant correlation between age and boundary width for a "rake"–"wake" identification task that differed from the present task only in the specific details of the stimulus continuum. The number of participants achieving ceiling-level performance was higher in this study than in the previous study, which could be interpreted as evidence that the continuum used here is less suitable for measuring fine-grained differences in perception. On the other hand, participants in this study had an older average age than participants from the McAllister and Tiede's study (median age of 13;2 vs. 11;6), which could account for the greater prevalence of ceiling-level performance and potentially also the lack of a statistically significant association with age. With respect to the discrimination task, we are aware of no previous research examining maturational changes in discrimination of the /ɹ/–/w/ contrast over the course of later children and adolescence. However, it has been suggested that phonetic discrimination tasks should have relatively high sensitivity to detect ongoing refinement in older participants. For example, in their study of perception of place of articulation in stimuli with synthetically manipulated transitions, Sussman and Carney (1989) found no effect of age on performance on an identification task, but they did find that discrimination ability was not yet adultlike at 10 years of age. Furthermore, in this study, there was no evidence of a ceiling effect that could account for the lack of a statistically significant association with age on the discrimination task. This is an area that warrants further research with a wider range of participant ages.

Somatosensory Measures

Our findings regarding associations between tasks assessing somatosensory acuity were broadly in line with our original hypotheses. Specifically, we predicted that we would observe stronger associations between the bite block and articulatory awareness tasks, which we view as primarily assessing the proprioceptive aspects of somatosensation, than between those tasks and the oral stereognosis task, which primarily measures tactile acuity. In addition, stereognosis is a nonspeech oral somatosensory task, while the bite block and the articulatory awareness task are speech specific. Given the various differences that exist between the use of the speech structures in speech versus nonspeech tasks such as chewing and swallowing (e.g., Kent, 2015), it is possible that somatosensory skills differ for speech and nonspeech tasks, further contributing to the lack of association between the oral stereognosis task and the other somatosensory tasks. On the other hand, consistent with our prediction, there was a statistically significant association between performance on the articulatory awareness and the bite block task. Although both tasks were designed to measure the sense of oral proprioception, they do so in very different ways—the bite block task requires implicit adaptation to the presence of a foreign object in the oral cavity, whereas the articulatory awareness task calls for explicit reflection on the relative position of the articulators. To the best of our knowledge, this is the first demonstration that explicit and implicit measures of oral proprioception correlate across individuals. We note that Gritsyk et al. (2021) administered the same somatosensory tasks to female adults and found no statistically significant association between any of the measures, in contrast with the present results observed with a child/adolescent sample. However, Gritsyk et al.'s study was a pilot study with a small sample size (n = 18), so their null result may be attributable to low experimental power.

This study found no statistically significant association between age and performance on the oral stereognosis task assessing the tactile component of somatosensory function. Previous literature has reported mixed findings regarding the development of oral tactile acuity over the course of childhood and adolescence. Several older studies (Gisel & Schwob, 1988; Kumin et al., 1984; McDonald & Aungst, 1967) found evidence of a developmental progression of tactile acuity characterized by increasing accuracy on oral form recognition tasks. However, recent studies by Appiani et al. (2020) and Lee et al. (2022) found no association between age and tactile acuity as assessed with Von Frey filament and/or grating orientation tasks. Lukasewycz and Mennella (2012) administered an oral letter identification task that is a close counterpart of the oral stereognosis task administered here. They did not find a statistically significant difference in oral letter identification between children aged 7–10 years and adults, consistent with the present finding of no association with age on oral stereognosis. As noted above, due to COVID-19 masking requirements, the stereognosis task was completed by a smaller number of participants (n = 75) than the other tasks. This raises the possibility that lower experimental power could have limited our ability to observe an association with age in the context of this task. However, the recent studies by Appiani et al. and Lee et al. using grating orientation tasks to assess oral tactile acuity found no significant association with age despite having very large sample sizes, which lowers the likelihood that sample size could account for the null result reported here.

This study did find a statistically significant association between age and performance on the bite block task, as well as a similar trend for the articulatory awareness task that did not survive correction for multiple comparisons. Since we regard these tasks as primarily assessing proprioceptive aspects of somatosensory function, we interpret this finding as evidence that the ability to access and interpret oral proprioceptive feedback continues to undergo refinement in later childhood and adolescence. Relatively little previous research has examined the maturation of oral proprioception in older children. In a study of children and adolescents aged roughly 6–18 years, Lohman et al. (2001) found a positive association between age and performance on a lingual awareness task similar to the articulatory awareness task administered here. However, their task assessed only consonant sounds produced with linguopalatal contact, and they characterized it as primarily a tactile measure. In light of the common patterning between their task, the present articulatory awareness task, and the bite block task administered here, we suggest that all three tasks tap the construct of proprioceptive awareness, which continues to undergo refinement in later stages of development. This relatively understudied topic warrants further investigation in future research.

Sensory Trade-Off

Previous research has suggested that acuity may be independent across auditory-perceptual and somatosensory domains (e.g., Ghosh et al., 2010) or that speakers may even exhibit sensory trade-offs, such that individuals who are highly responsive to auditory feedback are less responsive to somatosensory feedback or vice versa (Lametti et al., 2012; Madison & Fucci, 1971). In their study applying simultaneous auditory and somatosensory perturbations during speech production, Lametti et al. (2012) found evidence of a "sensory preference," such that adult speakers who compensated more for the auditory perturbation compensated less or not at all for the somatosensory compensation and vice versa. This study returned no evidence of such a trade-off when averaging participants' scores (expressed in terms of deviation from the median for their age group) across the three tasks in each domain. However, limitations of the individual tasks acknowledged elsewhere in this article (e.g., the ceiling effect in the identification task) may have contributed to this null result. It should also be noted that the tasks administered here were not equivalent across domains; as indicated previously, our auditory tasks were specific to the phoneme /ɹ/, whereas the somatosensory tasks were more general. This contrasts with the study by Lametti et al. (2012), where auditory feedback and somatosensory feedback were altered in the same speech context. The extent to which speakers rely on one domain versus the other is in part dependent on the speech target of interest; for example, the relative importance of somatosensory feedback is thought to be greater for consonant than for vowel sounds (Ghosh et al., 2010). It is thus possible that our failure to find support for the idea of a sensory trade-off may be related to the differing tasks used to probe the two domains. It would be desirable for future research to assess auditory-perceptual and somatosensory functions for /ɹ/ in comparable tasks, such as deriving auditory and somatosensory gain values based on performance in a task of compensation for auditory perturbation (as described in Kearney et al., 2020), to test if evidence of sensory trade-off can be obtained in this context. On the other hand, Madison and Fucci (1971) found evidence of a sensory trade-off, even though they used nonoverlapping tasks to assess the auditory and somatosensory domains. Thus, it remains possible that the battery of tasks used here could provide evidence of a sensory trade-off, particularly if administered in a context that yields a wider range of variation in individual performance (e.g., in speakers with RSSD).

Limitations

This study has several limitations. While efforts were made to recruit a sizable sample, the participants who responded to our recruitment efforts were not fully balanced with respect to sex (58% female) and were not distributed evenly across years of ages within the range of 9–15 years. We had relatively few participants below 10 years of age in either sex (in part because younger children were not eligible to be vaccinated for COVID-19 for a significant fraction of the study duration) and no female participants below 9;6. This sparse representation limits our confidence in our ability to characterize perceptual ability in participants at the lower end of our age range. As discussed above, the number of data points obtained was further reduced for the bite block and especially the oral stereognosis tasks due to masking requirements during COVID-19.

The fact that participants were not evenly distributed with respect to sex raises a question as to whether sex could, in principle, have been a significant mediator of performance in this study. There is limited evidence that the biological factor of sex influences individuals' perceptual acuity at a basic level. However, previous research on adults has suggested that performance in speech perception tasks is more accurate when there is a match in sex between the talker and the listener (Yoho et al., 2019). In addition, in a study of 59 children with RSSD aged 9–15 years, Cialdella et al. (2021) found that sex interacted with auditory-perceptual acuity to predict response to treatment. In the somatosensory domain, Kumin et al. (1984) found that female children scored significantly higher than male children on an oral stereognosis task (n = 168). However, a number of subsequent studies (e.g., Appiani et al., 2020; Gisel & Schwob, 1988; Lee et al., 2022; Lukasewycz & Mennella, 2012) have reported no sex-related differences in somatosensory acuity. Given the limited and conflicting nature of previous research, we did not go into this study with any hypotheses about sex-related differences in sensory performance. However, we examined this possibility in a post hoc fashion in linear regressions that predicted sensory performance based on the categorical variable of sex, as well as the continuous variable of age. No correction for multiple comparisons was undertaken given the exploratory nature of the research question. No significant sex-related differences were observed for any auditory-perceptual measure, for oral stereognosis, or for the articulatory awareness task. For the bite block task, there was a nonsignificant trend toward smaller ED (i.e., better performance) in male than female participants ($\beta = -0.13$, $SE = 0.07$, $p = .057$). The possibility of sex playing a role in sensory acuity warrants further investigation in future research, but the bulk of the evidence from this study suggests that male and female participants show similar performance on sensory tasks in the age range studied here.

The tasks used to assess auditory-perceptual and somatosensory acuity in this study have several clear limitations. The oral stereognosis task used here is particularly noteworthy in this regard. Oral letter recognition tasks have been criticized on the grounds that some letters are easier to recognize than others and that differences in cognitive ability or reading skill could confound the measurement of oral somatosensory acuity (Appiani et al., 2020; Lee et al., 2022). Visuospatial ability could also pose a confound for the letter recognition task administered here, since letters are presented with the top facing the back of the oral cavity and must be mentally rotated for identification. In future studies, we are interested in switching to a grating orientation task, which measures oral tactile acuity without any of these confounds. The existence of well-powered normative studies using a grating orientation task to assess oral tactile acuity (Appiani et al., 2020; Lee et al., 2022) represents a further rationale for switching to this task.

Another task that warrants further review is the "rake"–"wake" identification task, which was limited by a clear ceiling effect: Roughly 40% of all participants in this study achieved the minimum boundary width of zero. However, we are hesitant to discard this task, which in a slightly different version was found to differentiate between children with typical speech and those with RSSD (Cialdella et al., 2021) and to predict treatment response among children with RSSD (Benway et al., 2021; Preston et al., 2020). This suggests that the task can still be useful for clinical purposes, even if it is insufficiently sensitive to differentiate performance in older children developing typically.[6] This study also supports continued use of the category goodness judgment task, which showed a significant association with age in our sample of older children. In contrast, the discrimination task was not found to be more sensitive than the identification task (i.e., it also showed no statistically significant association with age or category goodness judgment), and it is a fairly demanding task in terms of sustained attention.

The need for sustained attention is the focus of a final point of limitation to consider for this study (Price & Moncrieff, 2021). All of the tasks reported here engage participants in somewhat unnatural experimental situations; accurate performance requires sustained attention and compliance with task instructions. Thus, scores reflect participants' ability to sustain attention in a challenging experimental setting as well as their raw perceptual ability, and associations between scores across tasks may be driven partly by this factor. Associations with age may be particularly susceptible to the moderator of capacity for sustained attention. Comparing individuals against age-matched peers, rather than the full sample of typically developing children, could be one way to partly correct for the influence of attentional maturity on the performance of laboratory-based sensory tasks. We propose that the reference values for older and younger children reported in Table 2 can be used for this purpose.

Future Directions

As laid out in the introduction, this study is part of a broader program of research that aims to refine treatment recommendations for children with RSSD. We adopt a personalized learning framework, which suggests that outcomes can be optimized when learners are matched to training paradigms tailored to their specific profile of strengths and needs. We also posit that auditory and somatosensory acuity are relevant parameters to measure when examining outcomes for the sensorimotor skill of speech production. We aim to use the reference data obtained from this study to characterize children with RSSD as typical or atypical with respect to auditory acuity, somatosensory acuity, or both in order to select a treatment that targets their sensory needs.

Specifically, this study is a companion to an ongoing randomized controlled trial comparing ultrasound biofeedback, visual–acoustic biofeedback, and motor-based non-biofeedback treatment (McAllister et al., 2020). Visual–acoustic biofeedback provides enhanced information about a speaker's success in matching an auditory–acoustic target but provides only indirect information about articulator placement, which speakers ordinarily experience through somatosensory channels. Conversely, ultrasound biofeedback provides direct information about articulator placement but corresponds less transparently to the information received through auditory channels. This has led us to predict that biofeedback treatment will be most effective when the enhanced feedback is aligned with the learner's most significant deficit area; that is, children with a primary deficit in auditory perception will derive the

[6]In general, it is important to recall that our ability to observe associations with age in this study may have been restricted by the relatively limited nature of variability in the performance of older speakers developing typically. The tasks reported here may still be valid for identifying atypical patterns of sensory performance in children with RSSD. We have been administering the same tasks reported here to children with RSSD as part of the evaluation for the related treatment study, and our preliminary results (with data collection still ongoing) suggest that robust between-groups differences are present for all measures, including those that did not show a significant association with age in the sample of typically developing speakers reported here.

greatest benefit from visual–acoustic biofeedback, whereas children with a primary somatosensory deficit will benefit more from ultrasound (for additional discussion, see Benway et al., 2021; Li et al., 2019). In the future, we hope that the findings from this line of research will help clinicians to assess sensory function in children with RSSD in a more principled way and to use the results of such assessment to pair clients with a treatment approach that will optimize learning outcomes. To support this application, materials and instructions for all of the measures administered here are shared for clinical and research use at https://osf.io/bkasm/. If administered in accordance with the protocols and inclusionary criteria described here, results of these measures can be compared against the reference values provided in Table 2. Out of the various measures investigated here, we suggest that the tasks most readily adaptable to the clinical setting (based on ease of administration and scoring) are the category goodness judgment task for auditory perception and the articulatory awareness task for somatosensory perception.

Conclusions

This study obtained three measures of auditory-perceptual function and three measures of oral somatosensory function in typically developing children/adolescents aged 9–15 years. The goal was to establish reference data that can be used as a point of comparison for individuals with RSSD, particularly those with residual deviations affecting American English rhotics. Spearman's rho was used to examine associations between tasks in a given domain, as well as the association of each task with age. We also tested for an inverse relationship between auditory-perceptual and somatosensory acuity, which would support the hypothesis of a trade-off between sensory domains. Statistically significant associations were observed between two tasks in the auditory-perceptual domain and two tasks in the somatosensory domain. In addition, significant associations with age were found for one auditory-perceptual and one somatosensory task. There was no evidence of a trade-off in acuity between auditory and somatosensory domains. Given the large sample size and the range of tasks administered, we are confident that we have obtained a broad, multidimensional characterization of auditory-perceptual and oral somatosensory function in typically developing children aged 9–15 years and that the resulting reference values can be used to characterize sensory profiles in children with RSSD. In the future, we hope to apply these reference data to make tailored treatment recommendations for a child with a specific sensory profile, such as ultrasound biofeedback for a child with a deficit in somatosensory perception. Ultimately, results from this research could help clinicians better characterize sensory function in children with RSSD, which could in turn lead to optimized treatment choices.

Author Contributions

Samantha A. Ayala: Project administration, Writing – original draft, Writing – review & editing, Visualization, Data curation, Formal analysis. **Amanda Eads:** Project administration, Writing – original draft, Writing – review & editing, Visualization, Data curation. **Heather Kabakoff:** Conceptualization, Project administration, Formal analysis, Visualization, Writing – review & editing **Michelle T. Swartz:** Project administration, Writing – review & editing. **Douglas M. Shiller:** Software, Methodology, Writing – review & editing. **Jennifer Hill:** Data curation, Formal analysis. **Elaine R. Hitchcock:** Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. **Jonathan L. Preston:** Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. **Tara McAllister:** Funding acquisition, Project administration, Writing – original draft, Writing – review & editing.

Data Availability Statement

The data sets generated during and/or analyzed during this study are available in the Open Science Framework repository, https://osf.io/2hy65/.

Acknowledgments

Research reported in this publication was supported by National Institute on Deafness and Other Communication Disorders Awards F31DC018197 (H. Kabakoff, PI), R01DC017476 (T. McAllister, PI), and R01DC017476-S1 (S. A. Ayala, trainee). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors express gratitude to José A. Ortiz and Nina R. Benway for the development of custom software used for stimulus presentation and data analysis. The authors also acknowledge the valuable contributions of the following study clinicians and research assistants: Megan Leece, Laura Ochs, Lily Jiao, Marcela Paz Lara, Leila Sandlin, and Josephine Oh.

References

Appiani, M., Rabitti, N. S., Methven, L., Cattaneo, C., & Laureati, M. (2020). Assessment of lingual tactile sensitivity in children and adults: Methodological suitability and challenges. *Foods, 9*(11), E1594. https://doi.org/10.3390/foods9111594

Attanasio, J. S. (1987). Relationships between oral sensory feedback skills and adaptation to delayed auditory feedback. *Journal of Communication Disorders, 20*(5), 391–402. https://doi.org/10.1016/0021-9924(87)90027-X

Audacity Team. (2019). *Audacity(R): Free audio editor and recorder* (Version 2.3.3) [Computer application]. https://audacityteam.org/

Baum, S. R., & Katz, W. F. (1988). Acoustic analysis of compensatory articulation in children. *The Journal of the Acoustical Society of America, 84*(5), 1662–1668. https://doi.org/10.1121/1.397181

Baum, S. R., & McFarland, D. H. (1997). The development of speech adaptation to an artificial palate. *The Journal of the Acoustical Society of America, 102*(4), 2353–2359. https://doi.org/10.1121/1.419619

Benway, N. R., Hitchcock, E. R., McAllister, T., Feeny, G. T., Hill, J., & Preston, J. L. (2021). Comparing biofeedback types for children with residual /ɹ/ errors in American English: A single-case randomization design. *American Journal of Speech-Language Pathology, 30*(4), 1819–1845. https://doi.org/10.1044/2021_AJSLP-20-00216

Berryman, L. J., Yau, J. M., & Hsiao, S. S. (2006). Representation of object size in the somatosensory system. *Journal of Neurophysiology, 96*(1), 27–39. https://doi.org/10.1152/jn.01190.2005

Bird, J., & Bishop, D. (1992). Perception and awareness of phonemes in phonologically impaired children. *European Journal of Disorders of Communication, 27*(4), 289–311. https://doi.org/10.3109/13682829209012042

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer (6.1.06)*.

Boyce, S. E. (2015). The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language, 36*(4), 257–270. https://doi.org/10.1055/s-0035-1562909

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics, 7*(3), 207–239. https://doi.org/10.1017/S0142716400007542

Cheng, H.-S., Niziolek, C. A., Buchwald, A., & McAllister, T. (2021). Examining the relationship between speech perception, production distinctness, and production variability. *Frontiers in Human Neuroscience, 15,* 660948. https://doi.org/10.3389/fnhum.2021.660948

Cialdella, L., Kabakoff, H., Preston, J. L., Dugan, S., Spencer, C., Boyce, S., Tiede, M., Whalen, D., & McAllister, T. (2021). Auditory-perceptual acuity in rhotic misarticulation: Baseline characteristics and treatment response. *Clinical Linguistics & Phonetics, 35*(1), 19–42. https://doi.org/10.1080/02699206.2020.1739749

Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America, 118*(3), 1661–1676. https://doi.org/10.1121/1.2000774

Crowe Hall, B. J. (1991). Attitudes of fourth and sixth graders toward peers with mild articulation disorders. *Language, Speech, and Hearing Services in Schools, 22*(1), 334–340. https://doi.org/10.1044/0161-1461.2201.334

DeJarnette, G. (1988). Formant frequencies (F1, F2) of jaw-free versus jaw-fixed vowels in normal and articulatory disordered children. *Perceptual and Motor Skills, 67*(3), 963–971. https://doi.org/10.2466/pms.1988.67.3.963

Dugan, S. H., Silbert, N., McAllister, T., Preston, J. L., Sotto, C., & Boyce, S. E. (2019). Modelling category goodness judgments in children with residual sound errors. *Clinical Linguistics & Phonetics, 33*(4), 295–315. https://doi.org/10.1080/02699206.2018.1477834

Edwards, J. (1992). Compensatory speech motor abilities in normal and phonologically disordered children. *Journal of Phonetics, 20*(2), 189–207. https://doi.org/10.1016/S0095-4470(19)30622-9

Engelen, L., van der Bilt, A., & Bosman, F. (2004). Relationship between oral sensitivity and masticatory performance. *Journal of Dental Research, 83*(5), 388–392. https://doi.org/10.1177/154405910408300507

Etter, N. M., Miller, O. M., & Ballard, K. J. (2017). Clinically available assessment measures for lingual and labial somatosensation in healthy adults: Normative data and test reliability. *American Journal of Speech-Language Pathology, 26*(3), 982–990. https://doi.org/10.1044/2017_AJSLP-16-0151

Feng, Y., Gracco, V. L., & Max, L. (2011). Integration of auditory and somatosensory error signals in the neural control of speech movements. *Journal of Neurophysiology, 106*(2), 667–679. https://doi.org/10.1152/jn.00638.2010

Flege, J. E., & Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica, 43*(4), 155–171. https://doi.org/10.1159/000261768

Flipsen, P. (2015). Emergence and prevalence of persistent and residual speech errors. *Seminars in Speech and Language, 36*(4), 217–223. https://doi.org/10.1055/s-0035-1562905

Franken, M. K., Acheson, D. J., McQueen, J. M., Eisner, F., & Hagoort, P. (2017). Individual variability as a window on production-perception interactions in speech motor control. *The Journal of the Acoustical Society of America, 142*(4), 2007–2018. https://doi.org/10.1121/1.5006899

Fucci, D. J. (1972). Oral vibrotactile sensation: An evaluation of normal and defective speakers. *Journal of Speech and Hearing Research, 15*(1), 179–184. https://doi.org/10.1044/jshr.1501.179

Fucci, D. J., & Robertson, J. H. (1971). "Functional" defective articulation: An oral sensory disturbance. *Perceptual and Motor Skills, 33*(3), 711–714. https://doi.org/10.2466/pms.1971.33.3.711

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics, 66*(3), 363–376. https://doi.org/10.3758/bf03194885

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Ménard, L., Guenther, F. H., Lane, H., & Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *The Journal of the Acoustical Society of America, 128*(5), 3079–3087. https://doi.org/10.1121/1.3493430

Gibson, A., & McPhearson, L. (1980). *Production of bite-block vowels by children*. Phonetic Experimental Research at the Institute of Linguistics (PERILUS) (Report II, pp. 26–41). University of Stockholm.

Gisel, E. G., & Schwob, H. (1988). Oral form discrimination in normal 5- to 8-year-old children: An adjunct to an eating assessment. *The Occupational Therapy Journal of Research, 8*(4), 195–209. https://doi.org/10.1177/153944928800800401

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation* (3rd ed.). Pearson.

Gritsyk, O., Kabakoff, H., Li, J. J., Ayala, S., Shiller, D. M., & McAllister, T. (2021). Toward an index of oral somatosensory acuity: Comparison of three measures in adults. *Perspectives of the ASHA Special Interest Groups, 6*(2), 500–512. https://doi.org/10.1044/2021_PERSP-20-00218

Guenther, F. H. (2016). *Neural control of speech*. MIT Press https://doi.org/10.7551/mitpress/10471.001.0001

**Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics, 95,* 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28*(4), 377–396. https://doi.org/10.1006/jpho.2000.0121

Hitchcock, E. R., Cabbage, K. L., Swartz, M. T., & Carrell, T. D. (2020). Measuring speech perception using the Wide-Range Acoustic Accuracy Scale: Preliminary findings. *Perspectives of the ASHA Special Interest Groups, 5*(4), 1098–1112. https://doi.org/10.1044/2020_PERSP-20-00037

Hitchcock, E. R., Harel, D., & McAllister Byun, T. M. (2015). Social, emotional, and academic impact of residual speech errors in school-aged children: A survey study. *Seminars in Speech and Language, 36*(4), 283–294. https://doi.org/10.1055/s-0035-1562911

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions: A developmental study. *Speech Communication, 53*(3), 417–430. https://doi.org/10.1016/j.specom.2010.11.005

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5,* 82. https://doi.org/10.3389/fnhum.2011.00082

Iverson, P., & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *The Journal of the Acoustical Society of America, 99*(2), 1130–1140. https://doi.org/10.1121/1.415234

Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S.-E., Chao, S.-C., Heller Murray, E. S., Scott, T. L., & Guenther, F. H. (2020). A simple 3-parameter model for examining adaptation in speech and voice production. *Frontiers in Psychology, 10,* 2995. https://doi.org/10.3389/fpsyg.2019.02995

Kendall, T., Thomas, E. R., & Kendall, M. T. (2018). *Package "vowels" in R* (Version 1.2.2). https://cran.rr-project.org/web/packages/vowels/vowels.pdf

Kent, R. D. (2015). Nonspeech oral movements and oral motor disorders: A narrative review. *American Journal of Speech-Language Pathology, 24*(4), 763–789. https://doi.org/10.1044/2015_AJSLP-14-0179

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics, 59,* 40–57. https://doi.org/10.1016/j.wocn.2016.08.006

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kumin, L. B., Saltysiak, E. B., Bell, K., Forget, K., Goodman, M. S., Goytisolo, M., Padden, J. V., Schroeter, N. C., & Thomas, S. (1984). Relationships of oral stereognostic ability to age and sex of children. *Perceptual and Motor Skills, 59*(1), 123–126. https://doi.org/10.2466/pms.1984.59.1.123

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *The Journal of Neuroscience, 32*(27), 9351–9358. https://doi.org/10.1523/JNEUROSCI.0404-12.2012

Lee, J., Russell, C. G., Mohebbi, M., & Keast, R. (2022). Grating orientation task: A screening tool for determination of oral tactile acuity in children. *Food Quality and Preference, 95,* 104365. https://doi.org/10.1016/j.foodqual.2021.104365

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Li, J. J., Ayala, S., Harel, D., Shiller, D. M., & McAllister, T. (2019). Individual predictors of response to biofeedback training for second-language production. *The Journal of the Acoustical Society of America, 146*(6), 4625–4643. https://doi.org/10.1121/1.5139423

Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics, 7*(2), 147–261. https://doi.org/10.1016/S0095-4470(19)31046-0

Lohman, P., & Fucci, D. (2000). Training in lingual awareness during production of isolated syllables of English. *Perceptual and Motor Skills, 91*(3, Suppl. 1), 1241–1246. https://doi.org/10.2466/pms.2000.91.3f.1241

Lohman, P., & Fucci, D. (2001). Comparison of tactile awareness in children during the production of isolated syllables of English. *Perceptual and Motor Skills, 92*(3), 699–705. https://doi.org/10.2466/pms.2001.92.3.699

Lohman, P., Fucci, D., & Marinellie, S. A. (2001). Tactile awareness in children and adults during the production of isolated syllables of English. *Perceptual and Motor Skills, 93*(3), 859–866. https://doi.org/10.2466/pms.2001.93.3.859

Lukasewycz, L. D., & Mennella, J. A. (2012). Lingual tactile acuity and food texture preferences among children and their mothers. *Food Quality and Preference, 26*(1), 58–66. https://doi.org/10.1016/j.foodqual.2012.03.007

Madison, C. L., & Fucci, D. J. (1971). Speech-sound discrimination and tactile-kinesthetic discrimination in reference to speech production. *Perceptual and Motor Skills, 33*(3), 831–838. https://doi.org/10.2466/pms.1971.33.3.831

McAllister Byun, T., & Tiede, M. (2017). Perception-production relations in later development of American English rhotics. *PLOS ONE, 12*(2), Article e0172022. https://doi.org/10.1371/journal.pone.0172022

McAllister, T., Preston, J. L., Hitchcock, E. R., & Hill, J. (2020). Protocol for Correcting Residual Errors With Spectral, Ultrasound, Traditional Speech Therapy Randomized Controlled Trial (C-RESULTS RCT). *BMC Pediatrics, 20*(1), 66. https://doi.org/10.1186/s12887-020-1941-5

McCormack, J., McLeod, S., McAllister, L., & Harrison, L. J. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology, 11*(2), 155–170. https://doi.org/10.1080/17549500802676859

McDonald, E. T., & Aungst, L. F. (1967). Studies in oral sensorimotor function. In J. F. Bosma (Ed.), *Symposium on oral sensation and perception* (pp. 202–220). Thomas.

Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., & Ferguson, M. A. (2011). Development of auditory processing in 6- to 11-yr-old children. *Ear and Hearing, 32*(3), 269–285. https://doi.org/10.1097/AUD.0b013e318201c468

Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: A timeline of development. *International Journal of Audiology, 46*(9), 460–478. https://doi.org/10.1080/14992020701383019

Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Current Biology, 16*(19), 1918–1923. https://doi.org/10.1016/j.cub.2006.07.069

Nasir, S. M., & Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature Neuroscience, 11*(10), 1217–1222. https://doi.org/10.1038/nn.2193

Newman, R. S. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. *The Journal of the Acoustical Society of America, 113*(5), 2850–2860. https://doi.org/10.1121/1.1567280

Parrell, B., & Houde, J. (2019). Modeling the role of sensory feedback in speech motor control and learning. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2963–2985. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0127

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25*(5), 382–407. https://doi.org/10.1016/j.jneuroling.2010.02.011

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America, 116*(4), 2338–2344. https://doi.org/10.1121/1.1787524

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., Stockmann, E., & Guenther, F. H. (2004). The distinctness of speakers' /s/–/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research, 47*(6), 1259–1269. https://doi.org/10.1044/1092-4388(2004/095)

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America, 130*(1), 461–472. https://doi.org/10.1121/1.3593366

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics, 13*(2), 253–260. https://doi.org/10.3758/BF03214136

Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory & Cognition, 3*(1), 7–18. https://doi.org/10.3758/BF03198202

Preston, J. L., Hitchcock, E. R., & Leece, M. C. (2020). Auditory perception and ultrasound biofeedback treatment outcomes for children with residual /ɹ/ distortions: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research, 63*(2), 444–455. https://doi.org/10.1044/2019_JSLHR-19-00060

Preston, J. L., Irwin, J. R., & Turcios, J. (2015). Perception of speech sounds in school-aged children with speech sound disorders. *Seminars in Speech and Language, 36*(4), 224–233. https://doi.org/10.1055/s-0035-1562906

Price, C. N., & Moncrieff, D. (2021). Defining the role of attention in hierarchical auditory processing. *Audiology Research, 11*(1), 112–128. https://doi.org/10.3390/audiolres11010012

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ruscello, D. M. (1995). Visual feedback in treatment of residual phonological disorders. *Journal of Communication Disorders, 28*(4), 279–302. https://doi.org/10.1044/10.1016/0021-9924(95)00058-x

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal, 8*(1), 289–317. https://doi.org/10.32614/RJ-2016-021

Smith, B. L., & McLean-Muse, A. (1987). Effects of rate and bite block manipulations on kinematic characteristics of children's speech. *The Journal of the Acoustical Society of America, 81*(3), 747–754. https://doi.org/10.1121/1.394843

Smith, D. J., Stepp, C., Guenther, F. H., & Kearney, E. (2020). Contributions of auditory and somatosensory feedback to vocal motor control. *Journal of Speech, Language, and Hearing Research, 63*(7), 2039–2053. https://doi.org/10.1044/2020_JSLHR-19-00296

Steele, C. M., Hill, L., Stokely, S., & Peladeau-Pigeon, M. (2014). Age and strength influences on lingual tactile acuity. *Journal of Texture Studies, 45*(4), 317–323. https://doi.org/10.1111/jtxs.12076

Sussman, J. E., & Carney, A. E. (1989). Effects of transition length on the perception of stop consonants by children and adults. *Journal of Speech and Hearing Research, 32*(1), 151–160. https://doi.org/10.1044/jshr.3201.151

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature, 423*(6942), 866–869. https://doi.org/10.1038/nature01710

Tyler, M. D. (2021). Perceived phonological overlap in second-language categories: The acquisition of English /r/ and /l/ by Japanese native listeners. *Language, 6*(1), 4. https://doi.org/10.3390/languages6010004

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122*(4), 2306–2319. https://doi.org/10.1121/1.2773966

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals–Fifth Edition (CELF-5)*. NCS Pearson.

Wohlert, A. B., & Smith, A. (1998). Spatiotemporal stability of lip movements in older adult speakers. *Journal of Speech, Language, and Hearing Research, 41*(1), 41–50. https://doi.org/10.1044/jslhr.4101.41

Wong, P. C. M., Vuong, L. C., & Liu, K. (2017). Personalized learning: From neurogenetics of behaviors to designing optimal language training. *Neuropsychologia, 98*, 192–200. https://doi.org/10.1016/j.neuropsychologia.2016.10.002

Yoho, S. E., Borrie, S. A., Barrett, T. S., & Whittaker, D. B. (2019). Are there sex effects for speech intelligibility in American English? Examining the influence of talker, listener, and methodology. *Attention, Perception, & Psychophysics, 81*(2), 558–570. https://doi.org/10.3758/s13414-018-1635-3

Zandipour, M., Perkell, J., Guenther, F., Tiede, M., Honda, K., & Murano, E. (2006). Speaking with a bite-block: Data and modeling. *Proceedings of the 7th International Seminar on Speech Production*, 361–368.