Auditory Acuity and Response to Altered Auditory Feedback in Children with and without Residual Speech Sound Disorder

Amanda Eads[1], Elaine Kearney[2], Asuka Koda[3], Elaine Hitchcock[4], Douglas Shiller[5], Frank Guenther[6], Tara McAllister[1]

[1]New York University, [2]Queensland University of Technology, [3]Bronx High School of Science, [4]Montclair State University, [5]University of Montreal, [6]Boston University
are326@nyu.edu; elaine.kearney@qut.edu.au; kodaa@bxscience.edu; hitchcocke@mail.montclair.edu; douglas.shiller@umontreal.ca; guenther@bu.edu; tkm214@nyu.edu

ABSTRACT

Children with residual speech sound disorder (RSSD) exhibit speech production patterns that differ from their dialect community and persist beyond the typical developmental window. Differences in auditory acuity or auditory feedback response may contribute to these persisting deviations. To assess this, the present study evaluated auditory feedback response in 31 children aged 9-15 with and without RSSD. Participants heard their speech in near real-time with the first formant frequency shifted upwards by 30%. We also administered an explicit behavioral measure of auditory-perceptual acuity for the phonetic contrast in the perturbation task.

The results indicate a trend toward greater compensation under altered feedback conditions by children with RSSD and a trend towards greater adaptation (i.e., persistence of learned changes after altered feedback was withdrawn) by children without RSSD. There were no significant differences in auditory-perceptual acuity between groups and no significant correlations between auditory-perceptual acuity and degree of compensation or adaptation.

Keywords: altered auditory feedback, sensory acuity, residual speech sound disorder, children

1. INTRODUCTION

Children with speech sound disorder (SSD) exhibit atypical speech patterns that negatively affect intelligibility, posing a barrier to participation in social and academic settings [1]. Delayed speech development typically resolves by 8-9 years old, but between 2-5% of speakers exhibit residual speech sound disorder (RSSD) that persists through adolescence or even adulthood [2].

The causes of these persisting speech deviations are still poorly understood, but it is known that typical speech production requires fine-grained representations of the sensory characteristics of different speech sounds, as well as the ability to update motor plans when sensory feedback indicates a mismatch with the intended target. Previous research shows that children with SSD/RSSD, on average, exhibit weaker speech perception than children with typical production [3, 4]. However, some children with SSD/RSSD perform in the average range on perceptual tasks, even for the contrast(s) they produce in an atypical fashion [3 -5]. It is possible that such children perceive the contrast between their productions and those of others in their environment, but they have difficulty updating their speech-motor routines in response to sensory feedback indicating a discrepancy.

This possibility was supported in previous research investigating compensation for perturbed auditory feedback in children with SSD aged 3-7 years [6]. In perturbation studies, specialized software [7] is used to transform the acoustics of participants' speech in near-real time, such that they hear their own voice with altered auditory-perceptual characteristics. On average, speakers tend to compensate for these real-time perturbations (i.e., they tend to shift their production in the opposing direction). Speakers also typically exhibit some degree of adaptation (i.e., changes in production that persist after the perturbation is removed), which is suggestive of updates to the feedforward plan for the speech sound in question. In a study of young children with and without SSD, Terband et al. [6] found that typically developing children were successful in compensating for perturbed auditory feedback affecting the vowel /e/, whereas children with SSD actually tended to follow the shift rather than compensate for it. This suggested that the children with SSD were successful in detecting the altered feedback but were unable to adjust their motor plans to offset the perturbation.

The present study examined both auditory-perceptual acuity and response to altered auditory feedback in an older population of children/ adolescents with and without RSSD. The children with RSSD specifically exhibited deviations affecting American English /ɹ/, a late-emerging sound that is commonly a target of speech intervention in later childhood. Based on previous literature, we hypothesized that we would observe statistically significant between-group differences in both auditory-perceptual acuity and compensation/ adaptation in response to perturbed feedback. However, we also expected to observe individual heterogeneity underlying these group-level effects. In the future, we hope that individual differences in performance on such tasks can be leveraged to make customized recommendations for treatment for children with RSSD (e.g., input-oriented intervention if auditory-perceptual acuity is atypical; ultrasound biofeedback intervention if participants have intact perceptual discrimination but have difficulty updating a speech-motor plan in response to sensory feedback). In addition, we investigated the association between auditory-perceptual acuity and compensation/adaptation for altered auditory feedback. Based on previous research [8], we hypothesized that individuals with more acute auditory perception would exhibit a stronger response to perturbation.

2. METHODS

Data collection occurred in two different university laboratory settings with approval from a single external Institutional Review Board (BRANY 18-10-393). Written informed assent and parent permission were obtained for all participants.

2.1. Participants

The data reported in this study come from 31 children aged 9-15 years old who live in New York or New Jersey and speak a rhotic dialect of American English in the home. The children are split into two groups. One group (n=12, 10 male, 2 female) has a history of RSSD and participated in a 16-week treatment study to aid in their development of the English /ɹ/ sound. The second group (n=19, 6 male, 13 female) has no history of RSSD. Parental reports confirm no history of speech-language-hearing difficulties, with the exception of speech production difficulty for the RSSD group. For inclusion, participants were required to pass a bilateral pure-tone hearing screening (20 dB HL at 500, 1000, 200, and 4000 Hz) administered on the day of testing.

2.2. Auditory-Perceptual Acuity Tasks

To assess the participants' auditory-perceptual acuity, an identification and discrimination task focused on the same phonetic context (/hɛd/ - /hæd/) were used in combination. Perceptual stimuli were presented on a laptop computer over Sennheiser HD429 headphones in a sound-shielded booth.

2.2.1 Auditory-Perceptual Identification

Participants completed a perceptual identification task implemented in Experigen [9]. STRAIGHT synthesis [10] was used to generate an 11-step continuum between naturally produced tokens of *head* and *had*, elicited from a typical young adult female talker from the same area as the participants [11]. Participants were instructed to identify each token by clicking the word *head* or *had* in a forced-choice task. After a practice phase featuring the continuum endpoints, participants heard each continuum step 8 times in random order (total 72 trials). A break was provided halfway through.

Performance on the task was analyzed by fitting a logistic function over the number of *head* responses for each step of the continuum. The phoneme boundary is defined as the 50th percentile of probability, where *head* and *had* responses are equally likely [12].

2.2.2 Auditory-Perceptual Discrimination

Participants subsequently completed a gamified AXB discrimination task implemented in MATLAB. A 240-step *head-had* continuum (a superset of the continuum used in the identification task) was used. The starting point was set in an individualized fashion based on the boundary location determined in the identification task.

In a gamified interface, participants heard a standard stimulus (X) and were instructed to identify which of two flanking stimuli (A and B) differed from the central stimulus. The distance between A and B began at half the continuum width and was adjusted in an adaptive staircase design that shifted from 8-down 4-up to 2-down 1-up, with step distance halving after every 3 reversals; the task terminated after 9 reversals. The mean distance across the last three reversals was treated as the just noticeable difference (JND), where a higher JND indicates lower perceptual acuity.

2.3. Altered Auditory Feedback Tasks

2.3.1 Task design and administration

Participants completed reflexive and adaptive altered auditory feedback paradigms during the session. This paper will only report on the adaptive paradigm.

In a protocol modified from Daliri et al. [13], during the adaptive paradigm, Audapter software [7] was used to shift the first formant frequency (F1) values of the target vowel /ɛ/ upwards 30% in near real time. Participants wore a headset microphone and over-ear headphones so they could hear their altered auditory signal in near real time (estimated 27 ms delay [14]). The stimuli included the words *bed*, *head*, and *Ted* along with representative pictures. Participants were asked to produce the word as soon as the picture appeared on the screen and to hold out the vowel sound until the picture disappeared. They were then given visual feedback regarding their production duration (“Too short”, “Just right”, or “Too long”). The window for “Just right” was 1100-1500 milliseconds. The stimuli were randomized in blocks of three so that the same stimuli would not be presented twice in a row.

The experimenter introduced the participants to the task during a practice session. The experimenter demonstrated the task expectations and elicited and gave feedback for 4 practice productions of “bed.” Participants then repeated “bed” 3 more times during a voice calibration phase, which allowed the experimenter to monitor microphone levels and select the best Linear Predictive Coding (LPC) setting for the participant. Individualized LPC settings were necessary given the developmental age range of our participants [15]. Finally, the adaptive paradigm included a baseline phase of 18 trials, an altered F1 phase of 36 trials, and an aftereffect phase of 18 trials where the perturbation was removed. During the altered F1 phase, the perturbation turned on at voice onset and remained on for the duration of the trial.

2.3.2 Measurement

The results of the altered auditory feedback task were measured in MATLAB using FLVoice software [16]. Two of the authors reviewed both the waveform and spectrogram for all 72 trials per participant to ensure the quality of productions and formant traces. Trials were excluded if they were less than 700ms in duration, included a false start/misproduction, or featured audio distortions in the microphone or headphone signal. LPC settings were adjusted if the initially selected setting yielded unstable formant tracking. A custom MATLAB script extracted the average F1 value in the 350-550ms window for each trial. This time window is expected to capture both within-trial compensation and trial-to-trial adaptation.

3. RESULTS

3.1. Auditory-Perceptual Acuity

The boxplots in Figure 1 represent JND in the auditory discrimination task for each group. The median for the group with RSSD is slightly lower, suggesting a slight trend toward more acute performance in this group of children. However, an independent-samples t-test revealed no significant difference between the groups ($t(28.92) = 0.85$, p = 0.4). The boxplots in Figure 1 show two outliers in the non-RSSD group, which could be contributing to the appearance of the slight trend.

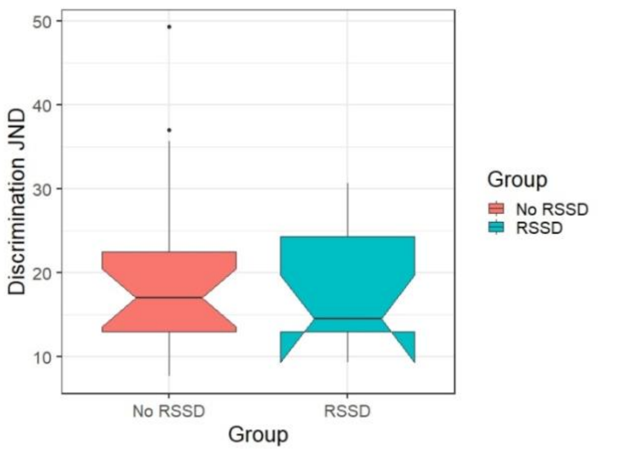

Figure 1: Just Noticeable Difference for each group in the auditory-perceptual discrimination task. Horizontal lines represent group medians, boxes represent the interquartile range, and outliers appear as points.

3.2. Altered Auditory Feedback Response

3.2.1 Altered F1 Phase

The boxplots in Figure 2 represent the average degree of compensation during the altered feedback phase for each group, relative to the baseline phase. The RSSD group has a lower median value than the non-RSSD group, indicating a greater degree of compensation during perturbation, and there is minimal overlap between the notches representing the confidence interval around the median for each group. However, an independent-samples t-test indicated that this difference was not statistically significant ($t(20.88) = 1.06$, $p = 0.3$).

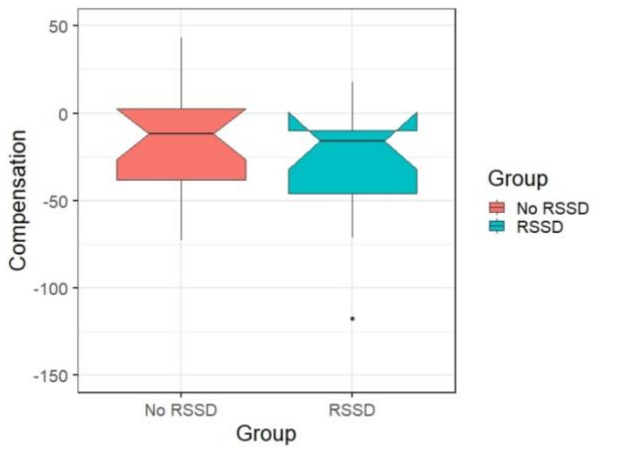

Figure 2: Percent compensation for each group during the altered F1 phase.

3.2.2 Aftereffect Phase

Figure 3 shows each group’s average degree of adaptation in the first three trials after the perturbation was removed, relative to the baseline phase. In contrast with the previous plot, Figure 3 shows a greater degree of adaptation for non-RSSD children than children with RSSD. One-sample t-tests were used to test if the degree of adaptation after the perturbation was withdrawn was significantly different from zero. The difference was significant for the group of children without RSSD ($t(18) = -2.72$, $p = .001$), but not for the group of children with RSSD ($t(11) = -1.04$, $p = .16$). However, the difference between groups was not statistically significant ($t(27.83) = -1.24$, $p = .22$).

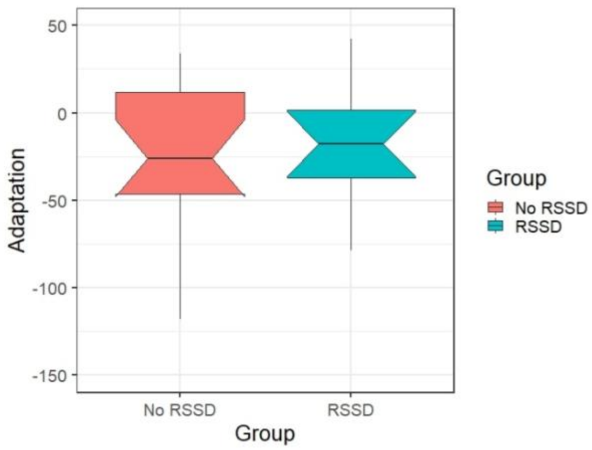

Figure 3: Percent adaptation for each group after withdrawal of the perturbation (aftereffect phase).

3.2. Auditory Acuity and Altered Auditory Feedback Response

Finally, we examined possible correlations between auditory acuity and the auditory feedback responses (percent compensation and adaptation). There was no significant correlation between discrimination JND and auditory feedback response either during the altered F1 phase ($r(29) = 0.18$, $p = 0.35$) or aftereffect phase ($r(29) = 0.05$, $p = 0.78$).

4. DISCUSSION

The results reported here are preliminary and must be interpreted with caution in light of a number of limitations. The number of participants is small, and the group of children with RSSD is roughly half the size of the group without RSSD. At the present time, the study is also limited in that the groups differed in sex breakdown, with a larger percentage of male participants in the RSSD group. We are continuing to recruit children with RSSD in connection with an ongoing clinical trial and will re-evaluate group differences after parity in group size is achieved.

In our preliminary sample, we found no significant difference in auditory acuity between the groups of participants with and without RSSD. While this was contrary to hypothesis, it was not entirely surprising, since we assessed perception using a sound contrast, /ɛ/-/æ/, that participants in the RSSD group did not have trouble articulating. Some previous literature has suggested that perceptual deficits in SSD may be specific to the contrasts that the speakers produce in error [17], although there is a lack of consensus on this point [3, 4].

In these preliminary data, children with RSSD showed a greater degree of compensation during perturbation than children without RSSD, although the difference was not statistically significant. This differs from a previous study of younger children [6], where a significant degree of compensation was observed in children with typical speech, but no compensation was observed in children with SSD. Given that the present study examined older children with a relatively mild speech distortion, it is not entirely surprising that our findings differed relative to [6]. Interestingly, whereas the RSSD group showed a numerically greater degree of compensation while the perturbation was in place, the group without RSSD showed a greater degree of adaptation after the perturbation was withdrawn. One-sample t-tests indicated that the magnitude of adaptation was significantly different from zero for the group without RSSD but not the group with RSSD, although the difference between the groups was not statistically significant. If this result is found to be robust when the complete sample is collected, it could constitute a suggestion that children with RSSD have more difficulty updating feedforward commands than children with typical speech production, and they may tend to rely correspondingly more on feedback in speech-motor control. This is broadly consistent with previous work on a different type of speech sound disorder, Childhood Apraxia of Speech [18].

5. CONCLUSION

While the current results only show group trends rather than statistically significant differences, they point to an interesting potential contrast between compensation during perturbation and adaptation after perturbation is withdrawn. Data collection is ongoing to achieve parity in group size and sex distribution. If the trends observed here are robust when the complete sample is collected, they will suggest a difference in feedback processing versus feedforward updating in children with and without RSSD. Additionally, SimpleDIVA modelling [19] is underway for this dataset, which will provide further insights into speech-motor control in this population.

6. REFERENCES

[1] J. McCormack, S. McLeod, L. McAllister, and L. J. Harrison. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. Int. J. Speech Lang. Pathol., 11(2): 155–170.

[2] Flipsen, P. (2015). Emergence and prevalence of persistent and residual speech errors. Semin. Speech Lang. 36(4):217–223.

[3] Cabbage, K., & Hitchcock, E. (2022). Clinical Considerations for Speech Perception in School-Age Children With Speech Sound Disorders: A Review of the Current Literature. Language, Speech, and Hearing Services in Schools 53(3):768-785.

[4] Hearnshaw, S., Baker, E., & Munro, N. (2019). Speech perception skills of children with speech sound disorders: A systematic review and meta-analysis. Journal of Speech, Language, and Hearing Research, 62(10): 3771-3789.

[5] Cialdella, L., Kabakoff, H., Preston, J. L., Dugan, S., Spencer, C., Boyce, S., Tiede, M., Whalen, D.H., & McAllister, T. (2020). Auditory-perceptual acuity in rhotic misarticulation: baseline characteristics and treatment response. Clinical Linguistics & Phonetics, 35(1): 19-42.

[6] H. Terband, F. van Brenk, and A. van Doornik-van der Zee. (2014).Auditory feedback perturbation in children with developmental speech sound disorders," J. Commun. Disord. 51:64–77.

[7] S. Cai, S. S. Ghosh, F. H. Guenther, and J. S. Perkell. (2010). Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization," J. Acoust. Soc. Am. 128(4):2033–2048.

[8] V. M. Villacorta, J. S. Perkell, and F. H. Guenther. (2007).Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception," J Acoust Soc Am. 122(4): 2306–19. doi: 10.1121/1.2773966.

[9] Becker, Michael and Jonathan Levine. (2013). Experigen – an online experiment platform. Available at http://becker.phonologist.org/experigen.

[10] H. Kawahara, M. Morise, H. Banno, and V. G. Skuk. (2013).Temporally variable multi-aspect N-way morphing based on interference-free speech representations: 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2013. doi: 10.1109/APSIPA.2013.6694355.

[11] A. Klaus, D. R. Lametti, D. M. Shiller, and T. McAllister. (2019). Can perceptual training alter the effect of visual biofeedback in speech-motor learning?. J. Acoust. Soc. Am., 145(2):805–817. doi: 10.1121/1.5089218.

[12] 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.

[13] Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S. E. (2018). Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. Developmental science, 21(2), e12521.

[14] Kim, K.S., Wang, H., and Max, L. (2020). It's about time: minimizing hardware and software latencies in speech research with real-time auditory feedback. Journal of Speech Language and Hearing Research, 63 (8): 2522-2534.

[15] Cheung, S. (2020). *Methodological Considerations and Findings of a Vowel Perturbation Study in Typically Developing Children* [Unpublished doctoral dissertation]. University of Toronto.

[16] Nieto-Castanon, A. (2022). GuentherLab FL Voice repository on GitHub. Retrieved from: https://github.com/GuentherLab/FLvoice.git

[17] Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. Journal of Speech and Hearing Disorders, 54(2), 193-208.

[18] Maas, E., Mailend, M. L., & Guenther, F. H. (2015). Feedforward and feedback control in apraxia of speech: Effects of noise masking on vowel production. Journal of Speech, Language, and Hearing Research, 58(2), 185-200.

[19] Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S.-E., Chao, S.-C., Murray, E. S. H., 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