



Improving Articulation *via* AI-Assisted Speech Practice in Children: A Randomized Controlled Pilot Trial

Shannan Conner*, Erin Kearney and Jessica Norris

Better Speech LLC, Online Speech Therapy Provider, Sanford, United States of America

***Correspondence:** Shannan Conner, Better Speech LLC, Online Speech Therapy Provider, Sanford, United States of America, E-mail: shannanc@betterspeech.com; DOI: <https://doi.org/10.56147/jbhs.3.4.164>

Citation: Conner S, Kearney E, Norris J (2026) Improving Articulation *via* AI-Assisted Speech Practice in Children: A Randomized Controlled Pilot Trial. J Biol & Heal Sci 3: 164.

Abstract

Background: Speech sound disorders are common in childhood, particularly for later-developing phonemes such as /s/ and /r/. Artificial Intelligence (AI)-assisted speech practice tools may increase treatment intensity and provide structured home practice, but controlled evidence in teletherapy contexts remains limited.

Objective: To evaluate whether integrating an AI-assisted speech practice tool with traditional teletherapy improves articulation accuracy of /s/ and /r/ sounds in children aged 6-17 years.

Methods: Twenty-three participants were randomized 1:1 to receive either traditional teletherapy plus AI-supported home practice (n=12) or traditional teletherapy alone (n=11). All participants completed eight clinician-led teletherapy sessions. Articulation accuracy was assessed at baseline and post-intervention using standardized word lists scored by licensed speech-language pathologists.

Results: Participants receiving AI-assisted practice demonstrated greater mean improvement (35.8%, SD 22.9) than those receiving teletherapy alone (27.5%, SD 22.4). The between-group difference was 8.4% points (95% CI -11.3-28.1), $t(21)=0.89$, $P=0.39$. The effect size was moderate (Hedges $g=0.36$). The pilot was underpowered to detect statistical significance.

Conclusions: Although not statistically significant, the direction and magnitude of the effect suggest that AI-assisted practice may enhance articulation outcomes when combined with teletherapy. Larger, prospectively registered trials with blinded scoring are warranted.

Keywords: Online speech therapy; Articulation disorders; Pediatric speech therapy; Telepractice; /r/ sound errors; /s/ sound errors

Received date: August 24, 2026; **Accepted date:** September 02, 2026; **Published date:** September 14, 2026

Introduction

Background

Speech Sound Disorders (SSDs) affect an estimated 3.6% of children at eight years of age and frequently involve later-developing phonemes such as /s/ and /r/, which require precise articulatory control [1]. Traditional speech-language therapy is effective for developmental

speech and language difficulties, but progress is influenced by treatment intensity, feedback quality and opportunities for structured practice outside therapy sessions [2-4]. Recent advances in Artificial Intelligence (AI) have enabled development of interactive speech practice tools capable of providing real-time feedback, automated error detection and high-frequency practice opportunities [5-7]. Despite their promise, controlled evidence evaluating the clinical impact of AI-assisted practice remains limited [8].



Gap in knowledge

Few randomized or controlled studies have examined whether AI-assisted practice enhances outcomes beyond traditional therapy alone, particularly for later-developing sounds. Evidence is especially sparse for school-aged children receiving teletherapy, despite research demonstrating that tele-practice yields outcomes comparable to in-person service delivery [9,10].

Rationale

Articulation learning depends heavily on repetition, feedback and distributed practice [3]. Integrating AI-supported home practice may strengthen treatment effects without increasing clinician burden. Technology-enhanced and biofeedback approaches have shown promise for residual /r/ errors, motivating evaluation of an AI-assisted adjunct within a real-world teletherapy workflow [11,12].

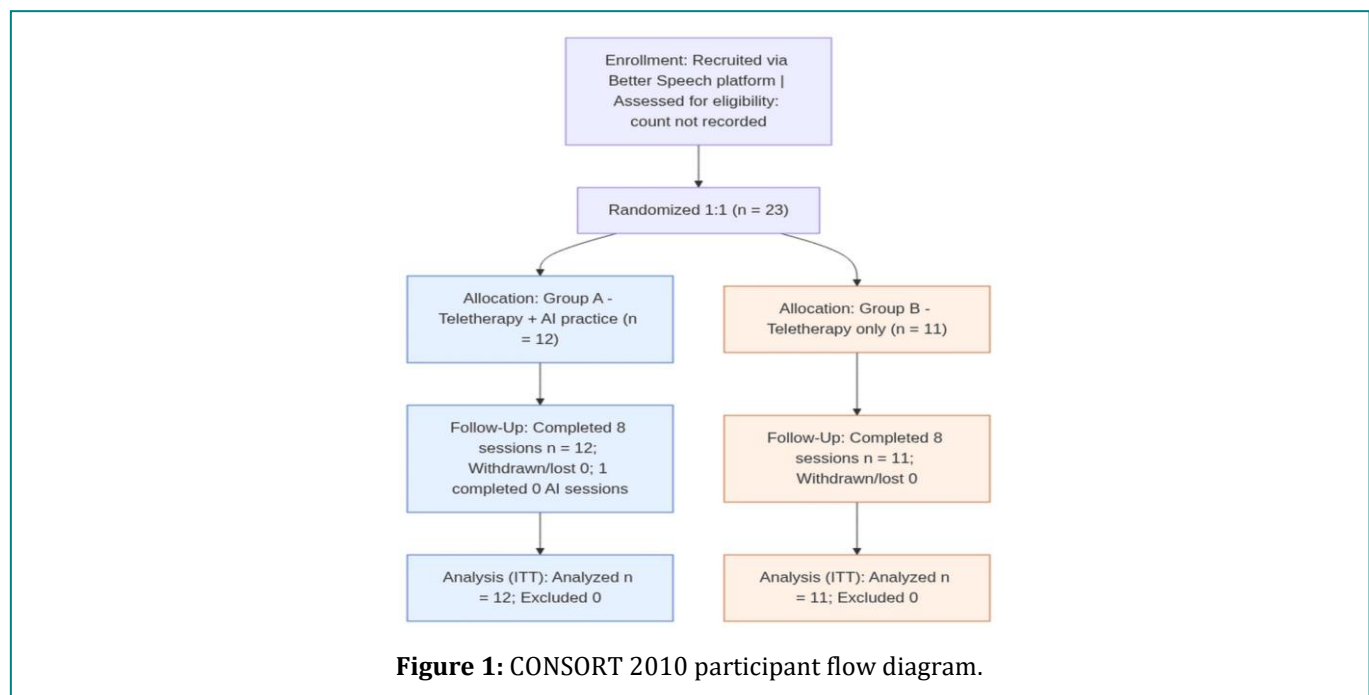
Objectives

The primary objective was to compare articulation improvement between children receiving traditional teletherapy alone and those receiving teletherapy plus AI-assisted practice. A secondary objective was to explore relationships between AI usage and articulation outcomes. We hypothesized that participants receiving AI-assisted practice would demonstrate greater improvement.

Methods

Study design

This study was a randomized controlled pilot trial designed to evaluate feasibility, estimate effect size and inform planning for a future definitive trial. The manuscript was prepared with reference to the CONSORT 2010 extension for pilot and feasibility trials. Participant flow is described narratively and depicted in the CONSORT diagram (Figure 1).



A target sample size of 20-30 participants was selected based on recommendations for pilot randomized controlled trials. The trial was conducted within an existing teletherapy service workflow. Because of its exploratory nature, the study was not prospectively registered; this is acknowledged as a limitation and future trials will be registered prior to enrollment.

Participants

Eligible participants were children aged 5-17 years who demonstrated word-level errors in /s/ or /r/ and resided in one of 20 eligible U.S. states. Exclusion criteria included documented developmental or neurological diagnoses (*e.g.*, ADHD, ASD) to reduce heterogeneity in

attention and learning profiles. The youngest enrolled participant was 6 years 6 months old, resulting in a final age range of 6;6 to 17;9 (years; months).

Recruitment and enrollment

Participants were recruited through the Better Speech online platform using email outreach, digital marketing and provider referrals. Eligibility was confirmed through screening and clinician evaluation. Caregivers provided informed consent prior to participation.

Randomization and allocation

Participants were randomized 1:1 to either (1)



traditional teletherapy plus AI-assisted practice (Group A) or (2) traditional teletherapy alone (Group B). Randomization was conducted within the clinical workflow; allocation concealment procedures were not implemented due to operational constraints of the teletherapy platform.

Intervention

Traditional teletherapy

All participants completed eight weekly clinician-led teletherapy sessions targeting /s/ or /r/ at the word level. Sessions were delivered by licensed speech-language pathologists using standardized articulation therapy techniques.

AI-assisted practice tool

Participants in Group A received access to Jessica, an AI-assisted articulation practice tool developed by Better Speech. Jessica provides automated feedback on phoneme production using acoustic pattern analysis and delivers corrective prompts in real time. The system supports high-frequency, repetition-based practice between clinician-led sessions.

Participants were instructed to use Jessica three times weekly for approximately 30 minutes per session. Actual usage varied widely (range 0-8 sessions; 0-174.5 minutes total) and is reported in **Table 1**. Raw audio was not stored; only anonymized performance metrics (*e.g.*, session count, practice duration, item-level accuracy) were transmitted.

Table 1: Participant characteristics and raw outcome data.

Participant #	Group	Age (yrs; mos)	Gender	Target sound	Baseline (S1)	Post (S8)	Improvement (pp)	AI sessions	Total AI time
62005	A	12;0	Female	/r/	1.00%	65.00%	64	8	91 min 34 s
62008	A	7;0	Female	/r/	50.00%	60.00%	10	5	64 min 26 s
62011	A	13;1	Female	/r/	5.00%	35.00%	30	3	53 min 38 s
62013	A	12;0	Female	/s/	50.00%	60.00%	10	4	45 min 2 s
62014	A	11;1	Male	/r/	5.00%	75.00%	70	5	53 min 1 s
62015	A	13;8	Male	/s/	30.00%	60.00%	30	5	152 min 3 s
62018	A	11;2	Male	/r/	50.00%	72.00%	22	4	61 min 23 s
62020	A	12;7	Male	/s/	40.00%	90.00%	50	0	0 min 0 s
62021	A	9;4	Female	/r/	23.00%	58.00%	35	3	46 min 17 s
62022	A	6;6	Male	/r/	0.00%	63.00%	63	2	44 min 10 s
62025	A	12;1	Female	/s/	37.00%	82.00%	45	4	174 min 30 s
62026	A	8;1	Male	/r/	97.00%	98.00%	1	3	93 min 6 s
62001	B	6;11	Male	/r/	32.00%	33.00%	1	NA	NA
62016	B	6;9	Female	/r/	34.00%	39.00%	5	NA	NA
62012	B	17;9	Female	/r/	5.00%	19.00%	14	NA	NA
62017	B	8;10	Male	/r/	5.00%	35.00%	30	NA	NA
62019	B	13;2	Male	/r/	7.00%	50.00%	43	NA	NA
62035	B	13;1	Female	/r/	0.00%	70.00%	70	NA	NA
62036	B	11;9	Female	/s/	34.00%	92.00%	58	NA	NA
62037	B	10;0	Female	/r/	6.00%	37.00%	31	NA	NA
62038	B	8;2	Male	/r/	3.00%	7.00%	4	NA	NA
62039	B	10;11	Female	/r/	31.00%	57.00%	26	NA	NA
62040	B	9;4	Male	/r/	70.00%	90.00%	20	NA	NA

Note: S1= Session 1 (initial session) S8= Session 8 (final session) pp = percentage points



Formal validation of automated scoring against human raters was not conducted as part of this pilot and is identified as a priority for future research.

Outcome measures

The primary outcome was articulation accuracy, calculated as the percentage of correctly produced target phonemes across standardized word lists. Assessments were conducted at baseline (Session 1) and post-intervention (Session 8). Only baseline and post-intervention measures were included in the analysis.

Blinding

Speech-language pathologists scoring the recorded word-list productions were not blinded to group assignment. Because recordings were collected within a clinical teletherapy workflow, blinding was not feasible. Future trials will incorporate blinded, independent raters.

Rater reliability

Inter-rater reliability was not formally assessed. Scoring procedures followed routine clinical practice. Future studies will include dual-rater scoring and reliability metrics.

Data privacy and security

All participant data were stored in compliance with HIPAA standards. Audio recordings were encrypted in transit and at rest.

Statistical analysis

Between-group differences in mean improvement were

evaluated using an independent-samples t test (two-tailed, $\alpha=0.05$). Effect size was calculated as Hedges g with the small-sample correction factor J. 95% confidence intervals were computed for both the mean difference and effect size.

Exploratory Pearson correlations examined relationships between AI usage (number of practice sessions; total practice time) and improvement within Group A. Analyses were conducted on an intention-to-treat basis; the single Group A participant who completed zero AI sessions was retained in the primary comparison.

All calculations were independently verified against the raw data presented in **Table 1**.

Results

Participant flow

Twenty-three participants were enrolled, randomized (Group A: n=12; Group B: n=11) and completed all study procedures. No participants withdrew or were lost to follow-up. One participant in Group A completed zero AI practice sessions but was retained in the intention-to-treat analysis. Participant flow is shown in the CONSORT diagram (**Figure 1**).

Baseline characteristics

Baseline demographic and clinical characteristics were broadly comparable between groups (**Table 2**). Mean age was similar (Group A: 10.7 years; Group B: 10.6 years). Baseline articulation accuracy was numerically higher in Group A (32.3%, SD 28.3) than in Group B (20.6%, SD 21.6), although this difference was not statistically significant ($t(21)=1.12$, $P=0.28$).

Table2: Baseline characteristics by group.

Characteristic	Group A (n=12)	Group B (n=11)	Test statistic	p-value
Age, mean (SD), (years; months)	10.7 (2.4)	10.6 (3.2)	t=0.08	0.93
Age range (years; months)	6;6-13;8	6;9-17;9	-	-
Female, n (%)	6 (50.0%)	6 (55.0%)	Fisher's exact OR=0.83	1
Male, n (%)	6 (50.0%)	5 (45.0%)	-	-
Target sound /s/, n (%)	4 (33.0%)	1 (9.0%)	-	-
Target sound /r/, n (%)	8 (67.0%)	10 (91.0%)	Fisher's exact OR=0.20	0.32
Baseline accuracy, mean SD, (%)	32.3 (28.3)	20.6 (21.6)	t=1.12	0.28
Note: Continuous variables were compared using independent-samples t-tests. Categorical variables were compared using Fisher's exact tests because of small cell counts. OR=Odds Ratio. n=number of participants				

The distribution of target sounds differed between groups: Group A included a higher proportion of /s/ targets (33%) compared with Group B (9%). Given the small sample size, these imbalances are acknowledged as limitations.

Primary outcome: Articulation improvement

Mean improvement in articulation accuracy from

baseline to post-intervention was:

- **Group A (teletherapy+AI practice):** 35.8% (SD 22.9).
- **Group B (teletherapy only):** 27.5% (SD 22.4).

The between-group mean difference was 8.4 percentage points (95% CI -11.3 to 28.1), $t(21)=0.89$, $P=0.39$. The effect size was Hedges g=0.36 (95% CI -0.44 to 1.15), representing approximately 30.5% greater



improvement in Group A relative to Group B.

As expected for a pilot trial, confidence intervals were wide and crossed zero, indicating that the study was underpowered to detect statistically significant differences.

Exploratory engagement findings (group A only)

Exploratory analyses examined associations between AI practice and improvement within Group A.

Across all 12 participants in Group A:

- Correlation between number of AI sessions and improvement: $r=0.09$.
- Correlation between total AI practice time and improvement: $r=-0.08$.

When restricting analyses to the 11 participants who completed at least one AI session:

- Correlation between number of sessions and improvement: $r=0.28$ (weak, nonsignificant).

These exploratory findings do not establish a dose-response relationship and should be interpreted cautiously given the small sample size and variability in engagement.

Safety

No adverse events were reported.

Participant-level data

Complete participant-level data, including baseline accuracy, post-intervention accuracy, improvement, AI usage and demographic characteristics, are presented in **Table 1**. These data enable independent reproduction of all statistical analyses.

Discussion

Principal findings

In this randomized controlled pilot trial, children receiving AI-assisted home practice demonstrated greater mean improvement in articulation accuracy than those receiving traditional teletherapy alone (35.8% vs. 27.5%). The absolute between-group difference of 8.4 percentage points and moderate effect size (Hedges $g=0.36$) suggest a potential benefit of integrating AI-supported practice into pediatric articulation therapy. Although the confidence intervals were wide and the results did not reach statistical significance, the direction and magnitude of the effect indicate that AI-assisted practice may enhance treatment intensity and support carryover between clinician-led sessions.

These findings align with established evidence that

increased practice opportunities and structured feedback contribute to improved speech-sound outcomes [4,13]. Prior meta-analytic work has demonstrated that speech-language interventions are effective for developmental speech and language disorders and that treatment dosage may influence outcomes [2]. The present study extends this literature by suggesting that an AI-supported home-practice tool may offer a scalable means of increasing treatment intensity without adding clinician burden.

The results also reflect principles of motor learning, which emphasize high-frequency, distributed practice and immediate feedback as mechanisms that facilitate skill acquisition and retention [3]. Jessica, the AI-assisted practice tool evaluated in this study, was designed to provide repeated opportunities for practice and real-time feedback outside therapy sessions. These characteristics may partially explain the numerically greater improvements observed in the intervention group.

Comparison with prior work

The findings are consistent with emerging evidence supporting technology-enhanced interventions for speech sound disorders. Computer-based and game-based speech therapy platforms have been shown to increase practice opportunities, improve engagement and support accessibility for children with SSDs [6,7]. Similarly, visual-acoustic and ultrasound biofeedback interventions have demonstrated positive outcomes for children with persistent residual /r/ distortions, leveraging augmented feedback to support motor learning [11,12].

More recently, Benway and Preston reported positive outcomes using artificial intelligence-assisted therapy for residual /l/ errors in a single-case experimental study [8]. Although the present study included both /r/ and /s/ targets and employed a randomized pilot design, the observed pattern of greater improvement in the AI-assisted group is broadly consistent with these preliminary findings and contributes additional evidence supporting the potential role of AI in speech intervention.

The study also adds to the growing literature demonstrating that tele-practice can produce outcomes comparable to traditional face-to-face service delivery [9]. Integrating AI-supported practice into a teletherapy model may further enhance treatment intensity and support carryover of newly acquired speech skills between sessions [14,15].

Interpretation

Although participants receiving AI-assisted practice demonstrated greater improvement, the absence of a clear dose-response relationship tempers causal interpretation. Engagement varied widely and improvement was not strongly correlated with either number of AI sessions or total practice time. This variability may reflect individual differences in learning readiness, baseline severity, family



support, adherence or responsiveness to feedback. As such, the findings should be interpreted as preliminary evidence of potential benefit rather than confirmation of efficacy.

Strengths

- Conducted in a real-world teletherapy environment
- Standardized assessment procedures
- Complete participant-level data enabling independent verification
- Intention-to-treat analysis
- Transparent reporting aligned with CONSORT pilot guidelines.

Limitations

This pilot trial was limited by its small sample size, lack of blinded scoring and variability in AI engagement. Scoring was performed by non-blinded clinicians and inter-rater reliability was not assessed. Baseline imbalances in articulation accuracy and target-sound distribution may have influenced results. Exclusion of participants with co-occurring diagnoses may limit generalizability. The trial was not prospectively registered. The AI tool's automated scoring was not validated against human raters in this study.

Power considerations

The study was not powered to detect statistically significant differences. The observed effect size ($g=0.36$) provides a preliminary estimate for planning a future definitive trial. Detecting a between-group effect of this magnitude with 80% power at $\alpha=0.05$ (two-tailed) would require approximately 120 participants per group; this estimate should be refined through formal power analysis.

Future directions

Future research should include larger, prospectively registered samples; blinded independent raters; inter-rater reliability metrics; validation of AI scoring against human raters; long-term follow-up and evaluation across additional speech sound disorders. Further work should also examine adherence, engagement and potential dose-response relationships.

Clinical implications

This pilot trial provides preliminary evidence that AI-assisted speech practice may enhance articulation outcomes when combined with traditional therapy. Structured, feedback-driven home-practice tools may help clinicians increase treatment intensity, improve carryover and support families seeking accessible practice opportunities outside therapy sessions. These implications remain tentative pending confirmation in adequately

powered trials.

Conclusion

Children who received AI-assisted speech practice demonstrated a 30.5% greater mean improvement in articulation accuracy compared with those receiving traditional teletherapy alone. Although the between-group difference did not reach statistical significance in this underpowered pilot trial, the direction and magnitude of the effect suggest that integrating AI-supported home practice may enhance articulation outcomes for school-aged children with speech sound disorders.

These findings provide preliminary evidence supporting the feasibility and potential benefit of AI-assisted practice as an adjunct to clinician-directed therapy delivered *via* tele-practice. Larger, prospectively registered randomized controlled trials with blinded scoring, validated AI metrics and adequate sample sizes are warranted to determine efficacy, explore dose-response relationships and evaluate generalizability across additional speech sound disorders.

Conflicts of Interest

The authors are employed by Better Speech LLC, which developed the AI-assisted practice tool evaluated in this study.

Data Availability

All data are available upon reasonable request.

Trial Registration

Not applicable. This exploratory pilot trial was not prospectively registered.

References

1. Wren Y, Miller LL, Peters TJ, Emond A, Roulstone S (2016) Prevalence and predictors of persistent speech sound disorder at eight years old: Findings from a population cohort study. *J Speech Lang Hear Res* 59: 647-673. [Crossref] [Google Scholar] [Indexed]
2. Law J, Garrett Z, Nye C (2004) The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *J Speech Lang Hear Res* 47: 924-943. [Crossref] [Google Scholar] [Indexed]
3. Maas E, Robin DA, Austermann HSN, Freedman SE, Wulf G, et al. (2008) Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang Pathol* 17: 277-298. [Crossref] [Google Scholar] [Indexed]
4. Wales D, Skinner L, Hayman M (2017) The efficacy of telehealth-delivered speech and language intervention for primary school-age children: A systematic review. *Int J Telerehabil* 9: 55-70. [Crossref] [Google Scholar] [Indexed]
5. Hair A, Monroe P, Ahmed B, Ballard KJ, Gutierrez-Osuna R (2018) Apraxia world: A speech therapy game for children



- with speech sound disorders. Proceedings of the 17th ACM Conference on Interaction Design and Children: 201-219. [Crossref] [Google Scholar]
6. Benway NR, Preston JL, Hitchcock E, Rose Y, Salekin A, et al. (2023) Reproducible speech research with the artificial intelligence-ready PERCEPT corpora. *J Speech Lang Hear Res* 66: 1986-2009. [Crossref] [Google Scholar] [Indexed]
 7. Furlong L, Erickson S, Morris ME (2017) Computer-based speech therapy for childhood speech sound disorders. *J Commun Disord* 68: 50-69. [Crossref] [Google Scholar] [Indexed]
 8. Benway NR, Preston JL (2024) Artificial intelligence-assisted speech therapy for /l/: A single-case experimental study. *Am J Speech Lang Pathol* 33: 2461-2486. [Crossref] [Google Scholar] [Indexed]
 9. Grogan-Johnson S, Schmidt AM, Schenker J, Alvares R, Rowan LE, et al. (2013) A comparison of speech sound intervention delivered by tele-practice and side-by-side service delivery models. *Commun Disord Q* 34: 210-220. [Crossref] [Google Scholar]
 10. Williams AL (2012) Intensity in phonological intervention: is there a prescribed amount? *Int J Speech Lang Pathol* 14: 456-461. [Crossref] [Google Scholar] [Indexed]
 11. Preston JL, Hitchcock ER, Leece MC (2020) Auditory perception and ultrasound biofeedback treatment outcomes for children with residual /r/ distortions: A randomized controlled trial. *J Speech Lang Hear Res* 63: 444-455. [Crossref] [Google Scholar] [Indexed]
 12. McAllister BT (2017) Efficacy of visual-acoustic biofeedback intervention for residual rhotic errors: A single-subject randomization study. *J Speech Lang Hear Res* 60: 1175-1193. [Crossref] [Google Scholar] [Indexed]
 13. Sugden E, Baker E, Munro N, Williams AL (2016) Involvement of parents in intervention for childhood speech sound disorders: A review of the evidence. *Int J Lang Commun Disord* 51: 597-625. [Crossref] [Google Scholar]
 14. Fagherazzi G, Goetzinger C, Rashid MA, Aguayo GA, Huiart L (2020) Digital health strategies to fight COVID-19 worldwide: Challenges, recommendations and a call for papers. *J Med Internet Res* 22: e19284. [Crossref] [Google Scholar] [Indexed]
 15. Preston JL, Leece MC, McNamara K, Maas E (2019) Intensive treatment with ultrasound visual feedback for speech sound errors in childhood apraxia. *Front Hum Neurosci* 13: 440. [Crossref] [Google Scholar] [Indexed]