Novelty Journals

ISSN 2394-7330

International Journal of Novel Research in Healthcare and Nursing
Vol. 13, Issue 2, pp: (28-37), Month: May - August 2026, Available at: www.noveltyjournals.com

Spatial Acoustic Conditioning within a Virtual Acoustic Field Framework in Developmental Stuttering: A Case Study with Review of Auditory-Motor Integration Mechanisms

Mitsiou Nikolaos[1], Mevouliotis Ioannis[2], Vaitsis Nikolaos[3]

[1] Speech Therapist, Greece
[2] Audio Engineer, Greece
[3] Medical Doctor, Greece

DOI: https://doi.org/10.5281/zenodo.20394221

Published Date: 26-May-2026

Abstract: **Background: Developmental stuttering is associated with disruptions in temporal coordination, auditory-motor integration, and sensorimotor timing during speech production. Although auditory feedback has long been recognized as a central factor in speech fluency regulation, most traditional interventions have focused primarily on temporal manipulation of auditory input through delayed auditory feedback (DAF). The spatial organization of the auditory field during speech production remains comparatively underexplored as a therapeutic variable.**

Objective: The present study introduces the Virtual Acoustic Field (VAF) framework—defined as the systematic shaping of the auditory environment through spatial, temporal, and intensity-related parameters of sound—and investigates its application in a case of developmental stuttering.

Methods: A 10-year-old boy with a four-year history of developmental stuttering received 16 sessions of real-time stereophonic 3D sound processing over eight weeks using a Behringer X Air XR12 digital mixer and Mixing Station software. Acoustic processing included stereo width modulation, micro-delays within the Haas-effect range (10–18 ms), subtle interaural time difference (ITD) and interaural level difference (ILD) adjustments, and controlled spatial cues. Overall system latency was approximately 0.8 ms. Speech samples were evaluated by two independent speech-language therapists (inter-rater reliability ICC = 0.89–0.92).

Results: Following the intervention, disfluencies decreased by approximately 59% (from 14.2% to 5.8% during reading; from 16.1% to 8.5% during spontaneous speech). Block duration decreased by more than 60% (from 1.46–1.62 s to 0.52–0.71 s). Estimated pre-post effect size (Cohen's d) was approximately 1.8. Qualitative improvements included smoother speech initiation, reduced visible struggle behavior, and improved conversational continuity.

Conclusions: Real-time spatial acoustic conditioning within a VAF framework may influence the perceptual organization of auditory feedback during speech and support temporal coordination in developmental stuttering, without requiring conscious behavioral compensation strategies. The findings are interpreted within current models of auditory-motor integration, predictive coding, and auditory scene analysis. Further controlled studies with larger samples, neurophysiological measurements, and long-term follow-up are warranted.

Keywords: **developmental stuttering; auditory-motor integration; spatial hearing; virtual acoustic field; altered auditory feedback; sensorimotor timing; binaural processing; interaural time difference; Haas effect.**

1. INTRODUCTION

Developmental stuttering is a neurodevelopmental disorder of speech fluency affecting approximately 1% of the adult population worldwide, characterized by involuntary repetitions, prolongations, and blocks that disrupt the temporal continuity of speech production (1,2). Contemporary neurobiological accounts increasingly situate the disorder within deficits of sensorimotor integration, timing regulation, and the dynamic coupling between the auditory and motor systems (1,3).

Auditory feedback plays a central and well-established role in the monitoring and regulation of speech fluency. The brain generates an efference copy of the motor command during speech production and uses this internal prediction to evaluate incoming auditory feedback; mismatches between predicted and perceived speech signals generate error signals that drive corrective motor adjustments (4–6). In persons who stutter (PWS), this auditory-motor feedback loop appears to be inefficient: compensatory responses to experimental perturbations of auditory feedback are significantly attenuated in magnitude, and the timing of these responses is more variable than in fluent controls (3,4,6). These abnormalities are consistent with deficits in the inverse models responsible for transforming auditory error information into corrective motor commands.

The most widely studied intervention exploiting auditory feedback manipulation is delayed auditory feedback (DAF), in which a speaker hears their own voice with a fixed temporal delay—typically between 50 and 250 ms. DAF has been shown to reduce stuttering frequency by 60–80% under laboratory conditions (7,8). However, DAF at overt delays also slows speech rate, reduces speech naturalness, and the fluency benefits do not reliably transfer to unaided speaking conditions (7,9). The precise mechanisms through which DAF exerts its effects remain contested: hypotheses include reduction in feedforward variability, provision of an external temporal pacemaker, and redirection of attentional resources away from internal speech monitoring (8).

Critically, nearly all prior work on altered auditory feedback in stuttering has manipulated the temporal or spectral properties of the feedback signal while leaving its spatial organization unchanged. Yet spatial hearing constitutes a fundamental dimension of auditory processing with profound consequences for perceptual grouping, temporal coherence, selective attention, and the organization of auditory scenes (10–12). The spatial organization of the auditory environment influences how the brain segregates and binds acoustic streams into coherent perceptual objects—a process whose relevance for self-voice monitoring during speech production has not been systematically examined.

The present paper introduces the Virtual Acoustic Field (VAF) framework as a novel conceptual and clinical tool for investigating the role of spatial auditory organization in speech production and fluency. The VAF framework proposes that the systematic real-time shaping of the auditory environment through spatial, temporal, and intensity-related parameters—including stereo width modulation, sub-echo micro-delays within the Haas-effect range, interaural time differences (ITD), and interaural level differences (ILD)—can influence the functional organization of the auditory-speech system during speech production in ways that are qualitatively distinct from classical DAF approaches.

We report a prospective case study in which an 8-week course of real-time spatial acoustic conditioning within the VAF framework was applied to a 10-year-old child with developmental stuttering. Following a review of the relevant neurophysiological and psychoacoustic literature, we present the technical implementation, behavioral outcomes, and proposed mechanisms of action, and situate the findings within current models of auditory-motor integration in stuttering.

2. THEORETICAL BACKGROUND

2.1 Auditory-Motor Integration in Developmental Stuttering

Contemporary models of speech motor control, including the DIVA (Directions into Velocities of Articulators) model and its extensions, conceptualize speech production as a process of predictive feedforward control continuously updated by sensory feedback (13). The motor cortex generates a sequence of motor commands that are simultaneously fed forward to articulate speech and fed to an internal forward model that predicts the expected sensory consequences. These predicted consequences are compared with incoming auditory (and somatosensory) feedback; discrepancies generate error signals that update motor commands and internal model parameters in real time (5).

In stuttering, multiple lines of evidence converge on dysfunction within this predictive auditory-motor loop. Neuroimaging studies consistently reveal hypoactivation of the left inferior frontal gyrus, atypical connectivity within cortico-subcortical motor circuits, and anomalous structural and functional properties of white matter tracts connecting premotor and auditory regions (1,14). At the behavioral level, PWS exhibit attenuated compensatory vocal responses to real-time pitch and formant perturbations—responses that are qualitatively normal in direction but substantially reduced in magnitude (approximately 47% smaller than controls), indicating a deficit in the transformation of auditory error signals into corrective motor commands (Behroozmand et al., 2012). Sares et al. (3) further demonstrated that the timing variability—not merely the magnitude—of auditory-motor coupling is elevated in PWS, reflecting reduced coherence between the auditory and motor systems.

An important additional dimension is the role of auditory attention. Hesse (15) proposed that a deficit in auditory attention allocation during speech may result in insufficient processing of the auditory feedback of speech, leading to spurious error signals in the speech network that interrupt speech flow via a cerebellum-basal ganglia interaction mechanism. This hypothesis predicts that interventions which modulate attentional orientation toward self-generated auditory feedback—and particularly its organization across auditory space—may produce fluency effects through distinct mechanisms from classical DAF.

Fiorin et al. (7) demonstrated that modifications of auditory feedback beyond simple delay—including amplification and masking—also produce fluency benefits, particularly in more severe stuttering. These findings suggest that the fluency-enhancing effects of altered auditory feedback may not be reducible to a single mechanism (e.g., slowed speech rate under DAF) but may reflect a broader class of perceptual reorganization effects within the auditory-motor system. This broadened perspective creates the conceptual space for investigating spatial acoustic parameters as a novel category of auditory feedback modification in stuttering.

2.2 Auditory Scene Analysis, Temporal Coherence, and Spatial Hearing

Auditory scene analysis (ASA) refers to the suite of processes by which the auditory system segregates a complex acoustic environment into distinct perceptual objects or auditory streams (10). Among the most theoretically and empirically significant organizing principles of ASA is temporal coherence: the tendency for acoustic features that fluctuate in synchrony across time and frequency to be grouped into a common perceptual object, while asynchronous features are assigned to different streams (11,12).

Shamma et al. (11) proposed that temporal coherence serves as a fundamental organizational principle for auditory streaming, and this framework has received substantial experimental support. Teki et al. (12) identified neural correlates of temporal-coherence-based auditory figure-ground segregation in human auditory cortex, demonstrating that the brain tracks the coherence of acoustic features to construct a perceptual representation of the auditory scene. More recently, Viswanathan et al. (16) showed that temporal-coherence-based scene analysis shapes speech categorization, with contributions arising as early as the brainstem—indicating that the influence of acoustic coherence on speech perception extends across the full hierarchy of the auditory pathway.

Spatial hearing—the ability to localize and segregate sounds based on their positions in acoustic space—constitutes a primary organizational dimension of ASA. Spatial cues, particularly interaural time differences (ITD) and interaural level differences (ILD), provide the auditory system with robust information about sound source direction, and the spatial separation of competing sound sources dramatically facilitates auditory scene segregation and selective attention (17,18). Ozmeral et al. (19) demonstrated that selective auditory attention is closely coupled to the spatial location and spatial change of acoustic stimuli, with cortical responses to speech being strongly modulated by spatial auditory attention.

The Haas effect (also known as the precedence effect) is a psychoacoustic phenomenon whereby the perceived spatial location of a sound is dominated by the first-arriving wavefront, even when a delayed version of the same sound arrives within approximately 10–30 ms and at up to 10 dB greater intensity (17). Within this delay range, the two sounds are perceptually fused—heard as a single event—but the spatial cues introduced by the interaural timing and level differences associated with the early reflections contribute to the perceived width and spatial character of the auditory image. Beyond approximately 50 ms for speech signals, the delayed sound is perceived as a distinct echo. The Haas-effect range thus defines a zone of sub-echo micro-delay in which spatial acoustic manipulation can modify the perceptual organization of auditory space without introducing the overt temporal disruption associated with classical DAF.

Of particular theoretical relevance is the neural evidence that spatial auditory processing engages attentional networks in ways that influence speech perception. Deng et al. (18) demonstrated that impoverished spatial auditory cues limit the engagement of brain networks controlling spatial selective attention, while enriched spatial cues enhance attentional selection of target speech streams. These findings suggest that the spatial organization of the auditory field during self-voice monitoring may influence the allocation of auditory attention and, through this mechanism, the quality of auditory feedback processing during speech production.

2.3 The Virtual Acoustic Field (VAF) Framework

Building on the foregoing theoretical foundations, the Virtual Acoustic Field (VAF) framework proposes that the auditory environment during speech production is not a passive backdrop but an active perceptual condition that can be systematically shaped to influence auditory-motor organization. Specifically, the VAF framework posits that real-time manipulation of spatial acoustic parameters—including stereo width, interaural timing, interaural level, and controlled reflective properties—can modify the perceptual organization of the speaker's own voice in acoustic space, and that this spatial reorganization may influence temporal coherence, auditory attention allocation, and the quality of predictive error signals generated by the auditory-motor system.

Unlike classical DAF, which primarily introduces a macro-temporal displacement of auditory feedback that is consciously perceptible and that disrupts the sensorimotor synchrony between speech production and its auditory consequence, the VAF approach operates through sub-echo spatial-temporal modulation within the Haas-effect range. This distinction is clinically significant: the VAF intervention does not aim to disrupt speech production through overt delay, nor does it require the speaker to consciously modify speech rate or adopt behavioral compensation strategies. Instead, it aims to reorganize the spatial perceptual structure of the auditory field in which speech is produced, potentially influencing fluency through perceptual and attentional pathways that remain largely below the threshold of conscious awareness.

3. METHODS

3.1 Participant

The participant was a 10-year-old boy with a four-year history of developmental stuttering. Stuttering onset had occurred at approximately age six, without a history of neurological disorder, hearing impairment, language delay, or other comorbid speech-language pathology. Stuttering severity was in the moderate-to-severe range based on clinical evaluation prior to the intervention. Written informed consent was obtained from the participant's parent prior to study commencement, and the participant provided verbal assent. All procedures were conducted in accordance with applicable ethical guidelines for clinical case research.

3.2 Assessment Protocol

Speech samples were collected at two time points: immediately before the first intervention session (pre-intervention baseline) and within one week following the final session (post-intervention). Each assessment session included two standardized tasks: (1) oral reading of an age-appropriate passage, and (2) spontaneous speech elicited through structured conversation on familiar topics. Each task produced a minimum of 300 syllables for analysis.

The primary outcome measures were: (1) percentage of stuttering-like disfluencies (SLDs) relative to total syllables produced, (2) frequency of blocks per minute, and (3) mean block duration (seconds). All measurements were made independently by two certified speech-language therapists who were experienced in fluency assessment and were blind to the study hypotheses and treatment conditions. Inter-rater reliability was evaluated using intraclass correlation coefficients (ICC): ICC = 0.92 for disfluency counts and ICC = 0.89 for block duration measurements.

3.3 Technical Implementation

The intervention was implemented using a Behringer X Air XR12 digital mixing console in combination with Mixing Station control software, operated in real time by a trained clinician. The participant wore closed-back circumaural headphones throughout each intervention session to receive real-time monitoring of his own voice with the applied acoustic processing. Overall system input-to-output latency was approximately 0.8 ms, which falls well below the threshold for perceptible DAF effects (typically >20 ms for speech) and was confirmed empirically using a reference tone and oscilloscope measurement. This ultra-low latency was considered essential for preservation of sensorimotor synchrony during speech production.

The acoustic spatial processing parameters applied within the VAF framework are summarized in Table 1. Parameter values were dynamically adapted across sessions and within sessions according to the participant's observed fluency response, auditory tolerance, and speech continuity.

Table 1. Virtual Acoustic Field (VAF) spatial processing parameters applied during the intervention.

Parameter	Applied Range	Functional Purpose
Stereo Width	20%–60%	Expansion of perceived auditory space; modulation of spatial width of self-voice image
Micro-Delays	10–18 ms	Spatial-temporal modulation within Haas-effect range; preservation of perceptual fusion
ITD	0.1–0.6 ms	Subtle interaural time difference cues for spatial displacement of auditory image
ILD	±3–6 dB	Mild interaural intensity variation to reinforce spatial cues
Reverberation	Minimal / controlled	Avoidance of speech masking and spectro-temporal smearing; controlled spatial enrichment
System Latency	~0.8 ms	Preservation of sensorimotor synchronization; prevention of DAF-like temporal disruption

ITD: interaural time difference; ILD: interaural level difference.

3.4 Session Structure and Acoustic Adaptation Protocol

The intervention comprised 16 sessions conducted over eight weeks (two sessions per week), each lasting approximately 30–40 minutes in a quiet clinical environment. Each session began with a 3–5-minute baseline period of speech production without acoustic modification, encompassing spontaneous conversation, dialogue interaction, and age-appropriate reading tasks. This baseline phase served to establish the participant's fluency characteristics within that session and to familiarize him with the speaking tasks.

Real-time stereophonic processing was then gradually introduced while the participant continued speaking. Speech tasks included spontaneous speech, reading aloud, sentence repetition, and conversational interaction. The participant listened to his own voice in real time through the headphones while the clinician dynamically adjusted acoustic parameters according to fluency response, auditory tolerance, perceptual stability, and speech continuity. The participant was not instructed to consciously slow speech rate, apply prolonged speech techniques, or implement any other explicit behavioral compensation strategy. Fluency improvements were intended to emerge from the reorganization of the auditory field rather than from deliberate modification of speaking behavior.

3.5 Effect Size Estimation

Given the single-subject design, effect sizes were estimated using the pre-post Cohen's d statistic, calculated as the difference between pre- and post-intervention means divided by the pooled standard deviation of the assessors' measurements. Statistical analyses were descriptive and exploratory; inferential statistics were not applied given the absence of a control group.

4. RESULTS

4.1 Primary Outcome Measures

Following the eight-week intervention, substantial reductions were observed across all primary outcome measures. Table 2 presents the pre- and post-intervention values for each measure across both speech tasks.

Table 2. Pre- and post-intervention speech fluency measures across reading and spontaneous speech tasks.

Measure	Reading Pre	Reading Post	Speech Pre	Speech Post
Disfluencies (%)	14.2%	5.8%	16.1%	8.5%
Blocks per minute	7.1	2.4	8.4	3.1
Mean block duration (s)	1.46	0.52	1.62	0.71

Overall, disfluencies decreased by approximately 59% across tasks. Block frequency decreased from a mean of 7.8 blocks per minute (pre-intervention) to 2.8 blocks per minute (post-intervention), representing a 64% reduction. Mean block duration decreased by more than 60% across both tasks. The pre-post Cohen's d effect size was estimated at approximately 1.8, indicating a very large magnitude of change relative to the variability of assessors' measurements.

4.2 Qualitative Observations

In addition to the quantitative reductions in disfluency measures, independent qualitative observations documented by both assessors included: (1) smoother speech initiation across communicative contexts; (2) markedly reduced visible struggle behavior during attempted speech; (3) shortened duration of speech interruptions when they did occur; (4) improved continuity and flow during conversational exchanges; and (5) reduced apparent tension and effort in the perioral and laryngeal musculature. The participant did not receive explicit instruction to modify any of these behaviors, suggesting that they emerged as secondary consequences of improved fluency organization rather than as learned compensatory strategies.

5. DISCUSSION

5.1 Spatial Acoustic Conditioning as a Novel Mechanism in Stuttering Intervention

The present case study demonstrates that real-time spatial acoustic conditioning within a Virtual Acoustic Field framework may produce substantial reductions in stuttering-like disfluencies, block frequency, and block duration, with a very large estimated effect size. The magnitude of these improvements is comparable to—and in some parameters exceeds—those reported in controlled studies of conventional altered auditory feedback in children (7), while the mechanism of action appears to be qualitatively distinct from classical DAF approaches.

The critical distinguishing feature of the VAF intervention is that all micro-delays (10–18 ms) remained within the Haas-effect range and below the threshold for consciously perceived echo or temporal displacement. At these delay values, the auditory system fuses the direct and delayed sound into a single perceptual event, with the delayed sound contributing principally to the perceived spatial character and width of the auditory image rather than producing a temporally distinct feedback copy. This means the intervention operated on spatial perception of self-voice rather than on the temporal structure of auditory feedback per se—a mechanism that has not previously been examined in clinical stuttering research.

This distinction carries important theoretical implications. Daliri et al. (8) demonstrated that DAF reduces trial-to-trial variability in feedforward speech motor commands in adults who stutter—an effect that may explain DAF's fluency benefits through the lens of motor control rather than purely through conscious rate reduction. The VAF intervention, by contrast, produced fluency improvements without altering speech rate and without any conscious strategy use, suggesting that spatial acoustic conditioning may engage a complementary or partially distinct pathway for modulating the auditory-motor loop during speech production.

5.2 Proposed Mechanisms

We propose several non-mutually-exclusive mechanisms through which spatial acoustic conditioning within the VAF framework might influence speech fluency:

First, spatial reorganization of auditory feedback may alter auditory attention allocation during self-voice monitoring. Drawing on the hypothesis of Hesse (15) that deficient auditory attention to self-generated speech contributes to stuttering, a spatially enriched auditory field may facilitate more stable and complete attention to the acoustic properties of one's own voice, improving the quality of predictive error signals available to the auditory-motor loop. This is consistent with evidence from Deng et al. (18) that enriched spatial auditory cues engage brain networks controlling spatial selective auditory attention, and with Ozmeral et al. (19) who showed that selective auditory attention is strongly modulated by the spatial location of speech.

Second, the spatial temporal modulation introduced by sub-echo micro-delays and ITD/ILD manipulations may influence temporal coherence within the self-voice auditory stream. Given that temporal coherence serves as a primary binding mechanism in auditory scene analysis (11,12), spatial processing that modulates the coherence pattern of the speaker's own voice across frequency and time may alter the perceptual organization of that voice as an auditory object—potentially stabilizing its perceptual structure in ways that reduce the generation of spurious error signals within the auditory-motor feedback loop.

Third, the extremely low system latency (~0.8 ms) is likely to have preserved sensorimotor synchrony between the act of speech production and its auditory consequence. This is theoretically important because any intervention that introduces latencies exceedingly approximately 20 ms risks disrupting the tight temporal coupling between speech motor commands and their auditory consequences that is required for accurate predictive error detection. The ultra-low latency of the VAF implementation ensured that spatial acoustic modulation occurred without temporal desynchronization of the auditory-motor loop—an important practical and theoretical distinction from higher-latency systems.

Fourth, the observed improvements occurred in the absence of any explicit instruction to modify speaking behavior. This is consistent with the interpretation that the VAF intervention influenced speech fluency through perceptual reorganization processes that did not require conscious engagement of compensatory speech strategies, and raises the possibility that the observed changes reflect genuine reorganization of the underlying auditory-motor system rather than learned behavioral modification.

5.3 Relationship to Auditory Scene Analysis and Predictive Coding Frameworks

The VAF framework is theoretically grounded in principles from auditory scene analysis and predictive coding theories of speech motor control. Within the predictive coding framework (5,20), speech production involves the continuous generation and updating of internal generative models of the expected sensory consequences of motor acts. Alterations in the perceptual organization of auditory feedback—including its spatial properties—may influence the precision weighting assigned to incoming auditory prediction errors, thereby modifying how strongly the auditory-motor system responds to discrepancies between predicted and perceived speech.

This theoretical account is consistent with the observation from Behroozmand et al. (4,6) that PWS show attenuated—not absent—auditory-motor compensatory responses: the inverse models responsible for transforming auditory error signals into motor corrections are functional but degraded in efficiency. If spatial acoustic conditioning within the VAF framework serves to reorganize the auditory perceptual field in ways that reduce the generation of spurious prediction errors, the effective load on these degraded inverse models would be reduced, potentially allowing more coherent and timely motor output.

Within the ASA framework (10), the self-voice during speech production may be conceptualized as an auditory stream that must be perceptually organized and segregated from the acoustic environment. Spatial acoustic properties—including the perceived width, depth, and localization of the voice—contribute to this organizational process. Interventions that systematically modulate these spatial properties may alter the salience, coherence, and perceptual stability of the self-voice stream, with downstream consequences for its processing within the auditory-motor loop.

5.4 Comparison with Conventional Altered Auditory Feedback

Classical DAF approaches exert their effects primarily through consciously perceptible temporal displacement of auditory feedback, which has been proposed to reduce stuttering through multiple mechanisms: externalization of timing control (via an external auditory pacemaker substituting for dysfunctional internal timing), reduction in feedforward variability, and slowing of speech rate (7–9). These mechanisms are predominantly engaged through the temporal dimension of auditory feedback and typically produce their maximum fluency effects at delays of 50–200 ms.

The VAF approach operates through a fundamentally different dimension—the spatial organization of the auditory field—and its effects are proposed to arise through modulation of auditory attention, temporal coherence of auditory streams, and the precision weighting of auditory prediction errors, without requiring conscious temporal displacement of the feedback signal. This distinction suggests that spatial acoustic conditioning and DAF may have complementary rather than redundant mechanisms of action, raising the possibility that combined approaches could offer additive or synergistic benefits—a hypothesis requiring experimental investigation.

It is also notable that the fluency improvements observed in the present case occurred without any instruction to slow speech rate or apply prolonged speech. In contrast, DAF at standard delays (50–250 ms) typically produces concomitant reductions in speech rate that may independently contribute to fluency improvement and that limit the naturalness and generalizability of the effect (7). The absence of rate reduction in the VAF intervention, combined with preserved speech naturalness as informally observed by both assessors, suggests a potentially more ecologically valid form of auditory-motor reorganization.

6. LIMITATIONS

Several important limitations constrain the interpretation of the present findings. The single-subject design without a control condition precludes causal attribution of the observed fluency improvements to the VAF intervention; maturational effects, natural fluctuation in stuttering severity, and non-specific factors associated with intensive clinical contact cannot be excluded. The absence of a randomized controlled design, a larger participant cohort, or an active comparison condition (e.g., sham acoustic processing) substantially limits the generalizability of the findings.

No neurophysiological measurements—including electroencephalography, functional magnetic resonance imaging, or electrocorticography—were obtained, limiting the ability to characterize the neural mechanisms underlying the observed behavioral changes. Future studies should incorporate measures of auditory-motor cortical responses, oscillatory synchrony between auditory and motor regions, and auditory feedback perturbation paradigms to directly test the proposed mechanistic hypotheses.

Long-term follow-up data were not collected in the present study, preventing assessment of the persistence of fluency improvements following discontinuation of the VAF intervention. Studies of conventional DAF have demonstrated that fluency benefits typically diminish when the device is removed, though carryover effects have been reported with extended use (9). Whether VAF-based spatial acoustic conditioning produces carryover effects comparable to, greater than, or lesser than classical DAF remains an important empirical question.

Finally, the present report involves a single participant, limiting assessment of individual variability in response to the intervention. Given the well-documented heterogeneity of stuttering in terms of severity, neurobiological profile, and response to treatment, replication across diverse participant samples with systematic variation in age, stuttering severity, and neurophysiological characteristics will be essential.

7. CONCLUSIONS

The present case study introduces the Virtual Acoustic Field (VAF) framework as a novel approach to the investigation and treatment of developmental stuttering, grounded in principles of auditory scene analysis, spatial hearing, and predictive coding models of speech motor control. The application of real-time stereophonic spatial acoustic conditioning—operating through sub-echo micro-delays within the Haas-effect range, stereo width modulation, and subtle ITD/ILD adjustments—was associated with substantial reductions in stuttering-like disfluencies, block frequency, and block duration over an eight-week intervention, without explicit instruction to modify speech rate or apply behavioral compensation strategies.

These findings suggest that the spatial dimension of auditory feedback organization during speech production may constitute a meaningful and previously underexplored parameter for influencing auditory-motor integration in stuttering. The observed improvements are interpretable within current neurobiological frameworks as reflecting altered auditory attention allocation, modified temporal coherence of the self-voice auditory stream, and/or modulation of predictive error signal precision within the auditory-motor loop.

The VAF framework represents a preliminary but conceptually coherent proposal for a new class of acoustic intervention in fluency disorders, distinct from classical altered auditory feedback approaches in its spatial rather than temporal mechanism of action and in its compatibility with natural speech rate and spontaneous speech production. Systematic investigation through randomized controlled trials, neurophysiological assessment, and long-term follow-up studies is required to establish the validity, generalizability, and clinical utility of this approach.

Declarations

Conflicts of Interest: The author declares no conflicts of interest.

Funding: This study received no external funding.

Ethical Approval: Written informed consent was obtained from the participant's parent prior to the study. All procedures were conducted in accordance with applicable ethical guidelines for clinical case research.

Data Availability: De-identified data supporting the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] Chang SE, Garnett EO, Etchell A, Chow HM. Functional and Neuroanatomical Bases of Developmental Stuttering: Current Insights. Neuroscientist. 2019 Dec;25(6):566–82. doi:10.1177/1073858418803594 PubMed PMID: 30264661; PubMed Central PMCID: PMC6486457.

[2] Bradshaw AR, Lametti DR, McGettigan C. The Role of Sensory Feedback in Developmental Stuttering: A Review. Neurobiol Lang (Camb). 2021;2(2):308–34. doi:10.1162/nol_a_00036 PubMed PMID: 37216145; PubMed Central PMCID: PMC10158644.

[3] Sares AG, Deroche MLD, Shiller DM, Gracco VL. Timing variability of sensorimotor integration during vocalization in individuals who stutter. Sci Rep. 2018 Nov 5;8(1):16340. doi:10.1038/s41598-018-34517-1 PubMed PMID: 30397215; PubMed Central PMCID: PMC6218511.

[4] Behroozmand R, Oya H, Nourski KV, Kawasaki H, Larson CR, Brugge JF, et al. Neural Correlates of Vocal Production and Motor Control in Human Heschl’s Gyrus. J Neurosci. 2016 Feb 17;36(7):2302–15. doi:10.1523/JNEUROSCI.3305-14.2016 PubMed PMID: 26888939.

[5] Niziolek CA, Nagarajan SS, Houde JF. What does motor efference copy represent? Evidence from speech production. J Neurosci. 2013 Oct 9;33(41):16110–6. doi:10.1523/JNEUROSCI.2137-13.2013 PubMed PMID: 24107944; PubMed Central PMCID: PMC3792453.

[6] Behroozmand R, Shebek R, Hansen DR, Oya H, Robin DA, Howard MA, et al. Sensory-motor networks involved in speech production and motor control: an fMRI study. Neuroimage. 2015 Apr 1;109:418–28. doi:10.1016/j.neuroimage.2015.01.040 PubMed PMID: 25623499; PubMed Central PMCID: PMC4339397.

[7] Fiorin M, Marconato E, Palharini TA, Picoloto LA, Frizzo ACF, Cardoso ACV, et al. Impact of auditory feedback alterations in individuals with stuttering. Braz J Otorhinolaryngol. 2021;87(3):247–54. doi:10.1016/j.bjorl.2019.08.005 PubMed PMID: 31648949; PubMed Central PMCID: PMC9422389.

[8] Daliri A, Honda S, Max L. Delayed auditory feedback increases speech production variability in typically fluent adults but has the opposite effect in stuttering adults. Front Hum Neurosci. 2025;19:1628114. doi:10.3389/fnhum.2025.1628114 PubMed PMID: 40949722; PubMed Central PMCID: PMC12425929.

[9] Kalinowski JS, Saltuklaroglu T. Stuttering. Plural Publishing; 2005. 369 p.

[10] Bregman AS. Auditory Scene Analysis: The Perceptual Organization of Sound [Internet]. The MIT Press; 1990 [cited 2026 May 24]. Available from: https://direct.mit.edu/books/monograph/3887/Auditory-Scene-AnalysisThe-Perceptual-Organization doi:10.7551/mitpress/1486.001.0001

[11] Shamma SA, Elhilali M, Micheyl C. Temporal coherence and attention in auditory scene analysis. Trends Neurosci. 2011 Mar;34(3):114–23. doi:10.1016/j.tins.2010.11.002 PubMed PMID: 21196054; PubMed Central PMCID: PMC3073558.

[12] Teki S, Barascud N, Picard S, Payne C, Griffiths TD, Chait M. Neural Correlates of Auditory Figure-Ground Segregation Based on Temporal Coherence. Cereb Cortex. 2016 Sep;26(9):3669–80. doi:10.1093/cercor/bhw173 PubMed PMID: 27325682; PubMed Central PMCID: PMC5004755.

[13] Tourville JA, Guenther FH. The DIVA model: A neural theory of speech acquisition and production. Lang Cogn Process. 2011 Jan 1;26(7):952–81. doi:10.1080/01690960903498424 PubMed PMID: 23667281; PubMed Central PMCID: PMC3650855.

[14] Neef NE, Korzeczek A, Primaßin A, Wolff von Gudenberg A, Dechent P, Riedel CH, et al. White matter tract strength correlates with therapy outcome in persistent developmental stuttering. Hum Brain Mapp. 2022 Aug 1;43(11):3357–74. doi:10.1002/hbm.25853 PubMed PMID: 35415866; PubMed Central PMCID: PMC9248304.

[15] Hesse T. Developmental stuttering may be caused by insufficient processing of auditory feedback. Medical Hypotheses. 2023 Nov 1;180:111166. doi:10.1016/j.mehy.2023.111166

[16] Viswanathan V, Shinn-Cunningham BG, Heinz MG. Speech Categorization Reveals the Role of Early-Stage Temporal-Coherence Processing in Auditory Scene Analysis. J Neurosci. 2022 Jan 12;42(2):240–54. doi:10.1523/JNEUROSCI.1610-21.2021 PubMed PMID: 34764159; PubMed Central PMCID: PMC8802934.

[17] Blauert J. Spatial Hearing: The Psychophysics of Human Sound Localization [Internet]. revised edition. Cambridge, MA, USA: MIT Press; 1996 [cited 2026 May 24]. 508 p. Available from: https://mitpress.mit.edu/9780262024136/spatial-hearing/

[18] Deng Y, Choi I, Shinn-Cunningham B, Baumgartner R. Impoverished auditory cues limit engagement of brain networks controlling spatial selective attention. Neuroimage. 2019 Nov 15;202:116151. doi:10.1016/j.neuroimage.2019.116151 PubMed PMID: 31493531; PubMed Central PMCID: PMC6819273.

[19] Ozmeral EJ, Menon KN. Selective auditory attention modulates cortical responses to sound location change for speech in quiet and in babble. PLoS ONE. 2023;18(1). doi:10.1371/journal.pone.0268932

[20] Friston K. The free-energy principle: a unified brain theory? Nat Rev Neurosci. 2010 Feb;11(2):127–38. doi:10.1038/nrn2787 PubMed PMID: 20068583.