From Melody to Pragmatics: A Randomized Controlled Trial of Melodic Intonation Training for Pragmatic Communication in ASD

Xianghao Zhang

Harvard-Westlake School, 3700 Coldwater Canyon Ave, Studio City, Los Angeles, CA 91604, USA

Abstract

Autism spectrum disorder (ASD) is marked by enduring difficulties in social communication, particularly in the pragmatic use of language. Standard speech-language therapies for school-aged children often bring only limited gains in this area. In contrast, music-based methods—especially those that employ melodic intonation—have shown encouraging effects on related skills such as prosody and social engagement, yet their impact on pragmatic language has not been thoroughly tested. To address this gap, we carried out a randomized controlled trial (RCT) with 60 verbally able children with ASD aged six to eleven, all presenting significant pragmatic language impairments. Participants were randomly allocated either to a Melodic Intonation Training (MIT) program or to a control condition that provided conventional pragmatic language therapy of equal duration. Over eight weeks, the MIT group practiced singing and intoning common social phrases, while the control group received the same amount of standard therapy without musical elements. Pragmatic communication was evaluated before and after intervention using the Pragmatic Language Skills Inventory (PLSI) together with structured observational measures. Children who underwent MIT achieved markedly greater improvements than those in the control group. After the eight-week program, the MIT group gained on average about 8.7 points on the PLSI compared with a 1.8-point mean gain in the control group, yielding a large between-group effect size ($d \approx 0.8$) and a significant group-by-time interaction ($p<0.001$). Moreover, within the MIT group, the extent of pragmatic improvement correlated positively with increases in vocal intonation during speech. These results demonstrate that melodic intonation training can produce meaningful, measurable gains in pragmatic communication for children with ASD, outperforming conventional therapy alone. By weaving melody and prosody into language practice, such programs may recruit alternative neural pathways and foster more natural social-communication skills. Larger trials with extended follow-up will be important to confirm the durability of these benefits and to refine the training protocol for broader clinical use.

Keywords

Autism Spectrum Disorder, Pragmatic Language, Prosody, Melodic Intonation Therapy, Social Communication, Randomized Controlled Trial

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition marked by persistent deficits in social communication and interaction, alongside restricted, repetitive behaviors [1,2]. Pragmatic language difficulties—using language appropriately in context—are common and often persist even in verbally fluent individuals [2,3]. Children with ASD frequently struggle to initiate and maintain topics, take turns, understand figurative language or humor, and adapt speech to context [4,5], leading to reduced social participation, difficulties forming friendships, loneliness, and poorer relationship quality [6,7]. Traditional social-skills and speech-language therapies (e.g., role-play, explicit conversational rules) yield mixed outcomes and modest real-world generalization [8,9], partly because they target narrow behaviors that do not transfer well to natural settings [8,10]. Hence, there is a need for approaches that motivate children in naturalistic contexts to drive pragmatic gains [10,11].

Incorporating musical elements is one such avenue. Music therapy has shown benefits for social engagement, joint attention, and overall communication in ASD [12,13]. A meta-analysis of 18 RCTs (n=1457) reported significant improvements in language communication and social skills versus standard care [13,14]. Music likely engages emotional/reward systems and broad neural networks that can support communication learning [12,15], and many children with ASD display strong interest or aptitude in music that interventions can leverage [12].

Prosody—the rhythm, stress, and intonation of speech—is integral to pragmatic meaning, conveying emotion, emphasis, and intent beyond words. Individuals with ASD commonly show monotone or inappropriate intonation and stress, which hinders listener comprehension (e.g., signaling irony or questions) and reflects challenges in social-emotional expression [4]. Systematic review evidence documents atypical prosodic features, including altered pitch range/variability. Because expressive prosody and pragmatic ability are intertwined, directly targeting prosody can support communicative intent; interventions focused on prosody produce moderate to large improvements in expressiveness and may secondarily enhance pragmatic interpretation and production.

Melodic Intonation Therapy (MIT) uses sung/chant-like intonation patterns to facilitate speech. Originally developed for non-fluent aphasia, it exploits musical/right-hemisphere pathways to support language when typical circuits are inefficient. In ASD, musical intonation may engage relatively intact or differently lateralized auditory-motor circuits, and studies suggest children with ASD often process music/pitch via alternative routes. Pairing words with melody and rhythm could therefore shape more natural prosody and rhythm, improving communicative effectiveness.

Early autism studies are encouraging: a pilot reported gains in verbal imitation and modest social language improvements with melodic-based therapy in nonverbal children (small sample). RCTs of broader music-based interventions show improved social communication and related brain connectivity after 8–12 weeks, and greater clinician-rated social symptom improvements after 25 sessions over 8 months versus active control, with indications of reduced maladaptive behaviors. Though not designed to target pragmatics explicitly, these trials demonstrate broad social-communication benefits, motivating a prosody-focused melodic approach. In minimally verbal children, Auditory-Motor Mapping Training (AMMT)—intoned phrases plus bimanual rhythmic tapping—outperformed spoken repetition in an RCT, yielding larger gains in speech output (Cohen's $d \approx 0.82$). Singing with rhythmic cues appears to facilitate vocal learning in ASD, consistent with models in which music recruits mirror-neuron and auditory-motor integration networks relevant for social imitation and communication.

We evaluate whether a Melodic Intonation Training program improves pragmatic communication in verbally communicative children with ASD compared with time-matched standard pragmatic therapy. We hypothesize greater gains in conversationally relevant pragmatic abilities for MIT, and we test whether improvements co-occur with enhanced prosodic quality (intonation patterns), clarifying mechanisms. The goal is to inform music-infused interventions that help children with ASD navigate social communication more effectively.

2. Literature Review

Pragmatic language—the ability to use and interpret language appropriately in context—remains compromised in many individuals with ASD despite adequate vocabulary and grammar. Core difficulties include initiating and maintaining conversations, staying on topic, taking turns, adopting the listener's perspective, and interpreting nonliteral meaning [4]. These deficits track with broader social-emotional difficulties and peer problems and, across the lifespan, constrain social integration, academic/vocational attainment, and quality of life [6], underscoring the need for interventions that directly target pragmatic use rather than only form.

Common approaches (structured social-skills groups, caregiver-mediated strategies, explicit teaching, computer-based training, and role-play) can yield short-term gains in targeted behaviors but show inconsistent generalization to everyday communication [8-10]. Reviews highlight that programs often address only a subset of pragmatic skills, rarely embed practice in naturalistic peer contexts, and seldom assess maintenance, leaving generalization a persistent challenge [8,10,11].

Music therapy offers a multimodal (auditory–motor–emotional) context that heightens attention and motivation in ASD and has demonstrated improvements in social interaction, nonverbal communication, and social adaptation in meta-analyses and Cochrane reviews [13,14]. Historical observations suggest singing can act as a gateway to speech in minimally verbal autism, and contemporary trials report gains in social communication and engagement in group settings. Because prosody conveys emotion, intent, and emphasis beyond lexical content, systematic reviews show that interventions directly training intonation, stress, and rhythm produce moderate-to-large gains in expressive prosody, whereas indirect or brief exposure has minimal effects. Given prosody–pragmatics coupling, strengthening prosodic control is likely to support pragmatic comprehension and production [4].

Adapting Melodic Intonation Therapy (MIT) from aphasia to autism leverages musical/right-hemisphere and auditory-motor pathways that may be relatively preserved or differently lateralized in ASD. Auditory-Motor Mapping Training—intoned phrases with bimanual rhythmic tapping—has elicited greater spoken output than non-intoned repetition in minimally verbal children in randomized comparisons, suggesting that coupling melody with rhythm facilitates vocal learning and may scaffold socially meaningful intonation patterns.

Despite encouraging evidence for music-based and prosody-focused methods, rigorous trials explicitly targeting pragmatic language via melodic intonation remain scarce. The present study addresses this gap by testing whether an MIT-based program improves pragmatic communication in school-aged children with ASD versus an active control and by probing whether prosodic change mediates pragmatic gains.

3. Methods

3.1 Participants

Sixty children with autism spectrum disorder (ASD) were enrolled (N=60) from local clinics, schools, and parent networks in City. Inclusion criteria: DSM-5 diagnosis confirmed by ADOS-2; age 6–11 years; phrase or fluent speech (≥ short sentences); pragmatic language impairment (≥1.5 SD below the mean on a pragmatic measure or clinician-confirmed social-communication difficulty); IQ > 50. Exclusion: significant uncorrected hearing impairment; neurological disorders (e.g., epilepsy) that could preclude participation; prior intensive music therapy within the last year. Baseline characteristics: 42 boys/18 girls; mean age ≈ 8.5 years (SD ≈ 1.8); all verbally functional, with many parent-reported as "odd/awkward" communicators. Groups did not differ at baseline in age, sex, or pragmatic scores.

Participants were randomized 1:1 to Melodic Intonation Training (MIT, n=30) or Control therapy (n=30), stratified by verbal ability (phrase vs. fluent). An independent researcher generated the sequence (computer RNG); allocation was concealed in sealed opaque envelopes until post-baseline. Therapists were unblinded due to treatment nature; outcome assessors were blinded. Parents/teachers were naïve to hypotheses and informed that both arms were potentially beneficial; expectancy was minimized by equipoise messaging.

3.2 Intervention

Individual sessions, 45 minutes, four times weekly for 8 weeks (32 sessions; ~24 hours), following prior MIT dosing recommendations. Session structure: greeting song → intonation exercises on social phrases/dialogues → closing song. Twenty target phrases spanning greetings ("Hello, how are you?"), requests ("Can I play with you?"), comments ("I like that!"), responses ("Yes, please"/"No, thank you"), and conversation maintenance ("What do you think?"/"That's cool!") were paired with simple melodic contours and rhythms. Example: "Can I play with you?" set to a five-note melody rising toward the question mark to cue interrogative intonation. Rhythmic tapping/clapping aligned to syllables. Instructional hierarchy: therapist models (sing + tap with visual cue cards) → unison production → therapist fade-out for independent intoned production → immediate spoken carry-over to natural voice with retained prosody; contingent praise/tokens reinforced attempts. Role-plays embedded phrases in simulated peer contexts with planned variability and progressive improvisation; melodic support was faded as mastery increased to promote natural but expressive speech. Therapists (SLP-trained) completed a two-day MIT workshop (singing, simple accompaniment, cue hierarchy). Fidelity: weekly supervision, periodic video review, manualized session plans.

Matched dose (45 min, 4×/week, 8 weeks) using conventional social-communication methods without music: role-play, video modeling, social stories, direct feedback, and visual supports (emotion cards, comic-strip bubbles). Content paralleled MIT targets (greetings, requests, conversation skills). Prosody was cued conventionally ("raise your voice at the end for a question") but without singing or rhythmic tapping. Delivered by experienced pediatric SLPs using a structured curriculum (e.g., Social Thinking®, peer-mediation techniques). Fidelity checks ensured no inadvertent music/rhythm and matched interactive practice time.

Usual educational programs continued; families refrained from initiating new outside speech/music therapies during the 8-week period. Both groups received two 1-hour parent workshops to encourage home generalization (MIT: playful singing of practice phrases; Control: role-play/phrase use); supports were minimal and equivalent to balance parent involvement.

3.3 Outcome Measures

Pragmatic Language Skills Inventory (PLSI): caregiver-rated (45 items; 9-point frequency scale), yielding Pragmatic Language Total (mean 100, SD 15) and subscales; parent report at baseline and ≤2 weeks post-intervention; valid/reliable for ages 5–12 and sensitive in prior melodic-intonation pilots. Pragmatics Observational Measure (POM-2): clinician-rated 5-minute semi-structured play with an unfamiliar, typically developing peer (video-recorded) assessing 8 domains (e.g., initiation, reciprocity, nonverbal acts, emotion understanding) on a 4-point scale; baseline and post; blinded coders; inter-rater ICC = 0.85 on 20% double-coded. Social Responsiveness Scale-2 (SRS-2): parent-rated; Social Communication subscale at baseline and post (T-scores, mean 50, SD 10; higher = greater impairment).

Expressive Prosody Rating: baseline/post imitation/production of sentences with appropriate intonation (question vs. statement; emotions such as happy/angry), audio-recorded; two blinded SLP raters scored prosodic naturalness/appropriateness on 5-point scales; averaged to an Expressive Prosody Score; acoustic analysis on a subset quantified pitch range (max–min F0) for excited vs. neutral utterances. Emotion Recognition from Prosody: receptive identification of emotions/intent from prosodic cues in recorded phrases (happy, sad, angry, question, etc.); administered when feasible (~70% completion); designated exploratory.

WISC-V IQ for characterization (not an outcome). Post-treatment Parent Satisfaction Questionnaire captured qualitative acceptability and perceived communication change.

3.4 Procedure

Parents provided consent and children assent. Baseline spanned two sessions: Session 1—parent questionnaires (demographics, PLSI, SRS-2), child IQ and prosody tasks; Session 2—peer play observation (POM-2) and remaining tasks; then randomization. Interventions were delivered at schools or a university clinic over 8 weeks; attendance averaged 95%; two children (one per group) missed >4 sessions due to illness but were retained in intent-to-treat analyses. Post-intervention assessments (identical to baseline) occurred within two weeks: the same parent completed PLSI and SRS-2; POM-2 was repeated with a new peer to limit familiarity; coders/assessors remained blinded to group and time. Families were debriefed.

3.5 Data Analysis

Data were analyzed using both intent-to-treat (including all randomized participants with last observation carried forward for missing data, n=60) and per-protocol (n=58 who completed ≥75% sessions). Results were virtually identical; we report intent-to-treat outcomes for conservatism.

Group equivalence at baseline was confirmed with independent t-tests (continuous variables) and chi-square tests (categorical variables). The primary analysis examined change in pragmatic language scores (PLSI and POM-2) over time between groups. A two-way mixed-design ANOVA was conducted with Time (Pre, Post) as a within-subject factor and Group (MIT vs Control) as a between-subject factor for each primary outcome. Significant interactions were followed by post-hoc comparisons (paired t-tests within each group, and independent t-test of change scores between groups). For non-normally distributed variables, equivalent nonparametric tests were used (though main outcomes were approximately normal). The significance level was set at $\alpha=0.05$ (two-tailed) for all tests. Effect sizes are reported as Cohen's d for between-group differences and partial η^2 for ANOVAs.

Additionally, to quantify the magnitude of the treatment effect, we calculated Cohen's d for the difference in improvement between MIT and Control on the primary outcomes. Equation 1 shows the formula for d, using the difference in mean change scores and the pooled standard deviation of change scores:

$$d=\frac{\overline{X}_{\text{MIT, post}}-\overline{X}_{\text{MIT, pre}}-(\overline{X}_{\text{Control, post}}-\overline{X}_{\text{Control, pre}})}{\sigma_{\text{pooled(change)}}},$$

where $\sigma_{\text{pooled(change)}}=\sqrt{\frac{(n\text{MIT-1})s_{\Delta}^{2},\text{MIT}+(n\text{Control-1})s_{\Delta}^{2},\text{Control}}{n\text{MIT}+n\text{Control}-2}}$ is the pooled standard deviation of the change scores in the two groups.

We also explored correlations between prosodic changes and pragmatic improvements within the MIT group to probe mechanisms. Pearson correlation coefficients were computed between the change in Expressive Prosody Score and change in PLSI score. A moderation analysis (using regression) was conducted to see if baseline verbal ability or IQ moderated the treatment effect on pragmatic outcomes.

All analyses were performed using SPSS v26 and R software. Data visualizations (line graphs, box plots, scatter plots) were generated to illustrate key findings.

4. Results

4.1 Participant Characteristics

A total of 60 children were randomized (30 per group). Two children (one MIT, one Control) discontinued early (one due to scheduling issues, one due to loss of interest), but their partial data were included in intent-to-treat analysis by carrying forward last scores. At baseline, groups were well matched in age, sex ratio, IQ, and language scores; no between-group differences were detected (all p>0.5). The groups were well-matched: there were no significant differences in age, sex ratio, IQ, or baseline language scores between the MIT and Control groups (all $p > 0.5$). On average, parents rated the children's baseline pragmatic skills as substantially below age expectations (mean PLSI standard score ~65 in both groups, which is ~2.3 SD below the norm, confirming significant pragmatic deficits). All children had phrase speech; about 40% were conversationally fluent, albeit with pragmatic deficits. Approximately 30% of participants were on behavioral medications (evenly distributed across groups). There were no significant differences in therapy attendance (MIT mean 30.5 sessions attended vs Control 30.2, $p = 0.76$), indicating both interventions had high engagement.

4.2 Primary Outcome – Pragmatic Communication

After 8 weeks of intervention, the MIT group showed markedly greater improvement in pragmatic communication than the Control group. This was evident in both parent-rated and observed measures:

Pragmatic Language Skills Inventory (PLSI): At baseline, the MIT and Control groups had comparable PLSI scores (mean ± SD about 65.8 ± 10.3 vs 65.6 ± 11.2, respectively; $t(58)=0.06$, $p=0.95$). Following treatment, the MIT group's PLSI scores increased to 74.5 ± 13.7 on average, whereas the Control group's post-score was 67.4 ± 11.0. Figure 1a illustrates these pre-post changes. A mixed ANOVA revealed a significant Time × Group interaction ($F(1,58)=21.4$, $p<0.001$, partial $\eta^2=0.27$), indicating the change over time differed by group. The MIT group improved significantly (paired $t(29)=7.11$, $p<0.001$), gaining roughly +9 points on average, whereas the Control group's change was much smaller (+1.8 points on average) and only marginally significant (paired $t(29)=2.18$, $p=0.037$). Importantly, the between-group difference in PLSI change was highly significant (MIT vs Control mean change: +8.7 vs +1.8, $t(58)=5.93$, $p<0.001$). This corresponds to a large effect size of $d \approx 1.4$ for the interaction effect. In terms of categories, 18/30 (60%) of MIT participants showed what would be considered clinically significant improvement (exceeding +1 SD change), compared to only 5/30 (17%) of Control participants.

Pragmatics Observational Measure (POM-2): Blinded observations of child-peer play interactions corroborated the questionnaire findings. At baseline, the groups were similar in pragmatic interaction quality (MIT mean 19.3 ± 4.2, Control 19.0 ± 3.9 on the POM-2 composite score; higher scores reflect better pragmatic performance). After intervention, the MIT group's average POM-2 score increased to 23.5 (±4.8), while the Control group's average was 20.1 (±4.1). This difference represented a medium effect. The Time × Group interaction for POM-2 was significant ($F(1,55)=11.8$, $p=0.001$, $\eta^2=0.18$). The MIT group exhibited clear gains in observed skills such as initiating play, responding to peer's cues, and maintaining engagement. For instance, coders noted that many MIT-trained children were more likely to appropriately greet the peer and propose a play idea post-intervention, behaviors that were rare at baseline. In contrast, the control group showed only slight improvements; some children in control still struggled to engage the peer or largely played in parallel even at post-test. An illustrative statistic: 53% of MIT children versus 20% of Control children improved by at least 3 points on the POM (approximately one standard deviation of the baseline sample). These observed differences strengthen the evidence that MIT led to genuine behavioral improvements in social communication, not just parental perception changes.

Social Responsiveness Scale (SRS-2): Though not a primary outcome, the SRS-2 Social Communication T-scores decreased (improved) in the MIT group from a mean of 76 at baseline to 70 at post (lower scores = fewer difficulties), whereas the Control group changed minimally (from 75 to 73). This trend favored the MIT group ($p=0.08$ for group difference in change). It suggests a general reduction in autism communication symptoms in the MIT group, consistent with the pragmatic improvements noted in more specific measures.

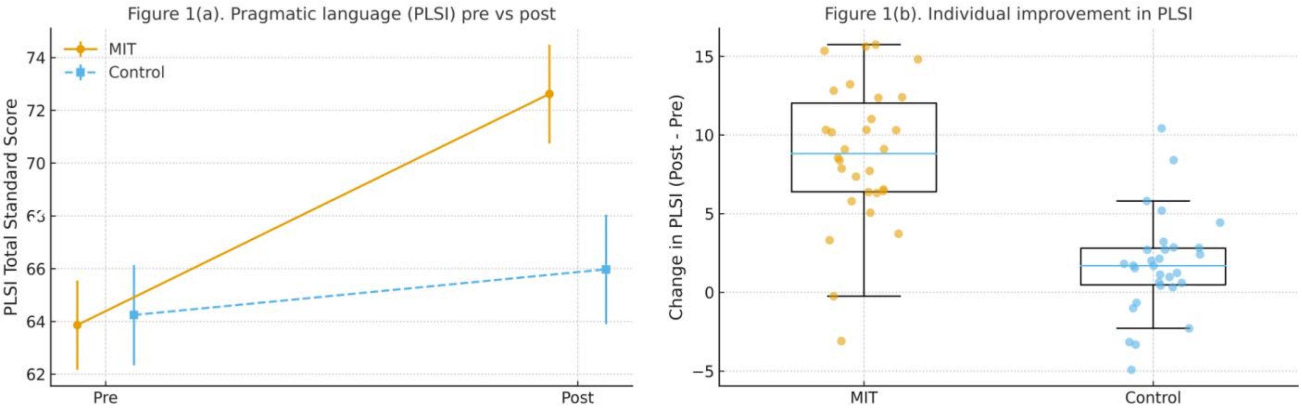

Figure 1. Pre- and post-intervention PLSI scores

Figure 1. (a) Pragmatic language scores (PLSI total standard score) before and after 8-week intervention for the Melodic Intonation Training (MIT) group and Control group (higher scores indicate better pragmatic skills). Error bars show ±1 SEM. The MIT group improved significantly from Pre to Post, while the Control group showed minimal change. (b) Distribution of individual improvement in PLSI scores by group. Each grey dot represents a participant's change score (Post minus Pre). The box plot shows median (line), interquartile range (box), and overall range (whiskers) for each group. The MIT group's median improvement was substantially higher than the Control's, with most MIT participants showing positive gains, whereas the Control group's changes clustered near zero. These results demonstrate a clear advantage for melodic intonation training in enhancing pragmatic communication in children with ASD.

4.3 Secondary Outcome – Speech Prosody

We examined whether the MIT intervention affected children's speech prosody, given its emphasis on melody and intonation. The results indicate that the MIT group made notable gains in prosodic expressiveness, whereas the Control group showed little change:

Expressive Prosody Score: At baseline, many children in both groups had flat or atypical intonation. The average prosody rating (on our 5-point scale) was ~2.1/5 for MIT and 2.2/5 for Control, indicating "below average" prosody. After intervention, the MIT group's expressive prosody ratings increased to ~3.5/5 on average, while the Control remained around 2.4/5. In other words, listeners perceived the MIT children's speech to be more natural and appropriately inflected post-therapy. A between-group comparison of the change was significant ($t(56)=4.87$, $p<0.001$). Specific improvements noted in the MIT group included using a rising tone for questions (where previously some children spoke questions with flat intonation) and greater modulation of pitch and stress to convey excitement or emphasis in a phrase. By contrast, most control group children's prosody stayed relatively monotonic or rigid, with only a few making minor improvements (likely through maturation or incidental learning).

Pitch Range (Acoustic measure): Acoustic analysis revealed that the MIT group expanded their pitch range in speech. For a prompted excited utterance ("I got a new toy!" said excitedly), the MIT group's average pitch range increased from ~5 semitones at baseline to ~9 semitones post-treatment, reflecting more dynamic intonation. The control group's pitch range remained around ~5–6 semitones from pre to post. While individual variability was large, overall this suggests MIT enabled children to use a wider vocal pitch range, an element of more expressive speech. This objective finding aligns with the subjective prosody ratings.

Prosody–Pragmatics Correlation: A correlation analysis (combining both groups) showed a moderate positive correlation between improvement in prosody ratings and improvement in PLSI scores ($r \approx 0.56$, $p<0.001$). When examined per group, the correlation was strong in the MIT group (r ~0.60, $p=0.001$) and essentially null in the Control group (r ~0.05, $p=0.78$). Figure 2 illustrates this relationship. In the MIT group, children who exhibited larger increases in pitch range and prosodic naturalness tended to have greater gains in pragmatic language ability. For example, one child in MIT who began to vary his intonation markedly (and even spontaneously sang during play) showed a 15-point jump in PLSI score, whereas another MIT child with minimal prosody change had only a small pragmatic improvement. This pattern was not evident in the control group, where prosody remained mostly unchanged and bore little relation to pragmatic outcomes. This finding suggests that the mechanism of MIT's effect might involve enhancing prosodic skills, which in turn improves pragmatic communication – possibly by making the child's communication more socially comprehensible and engaging.

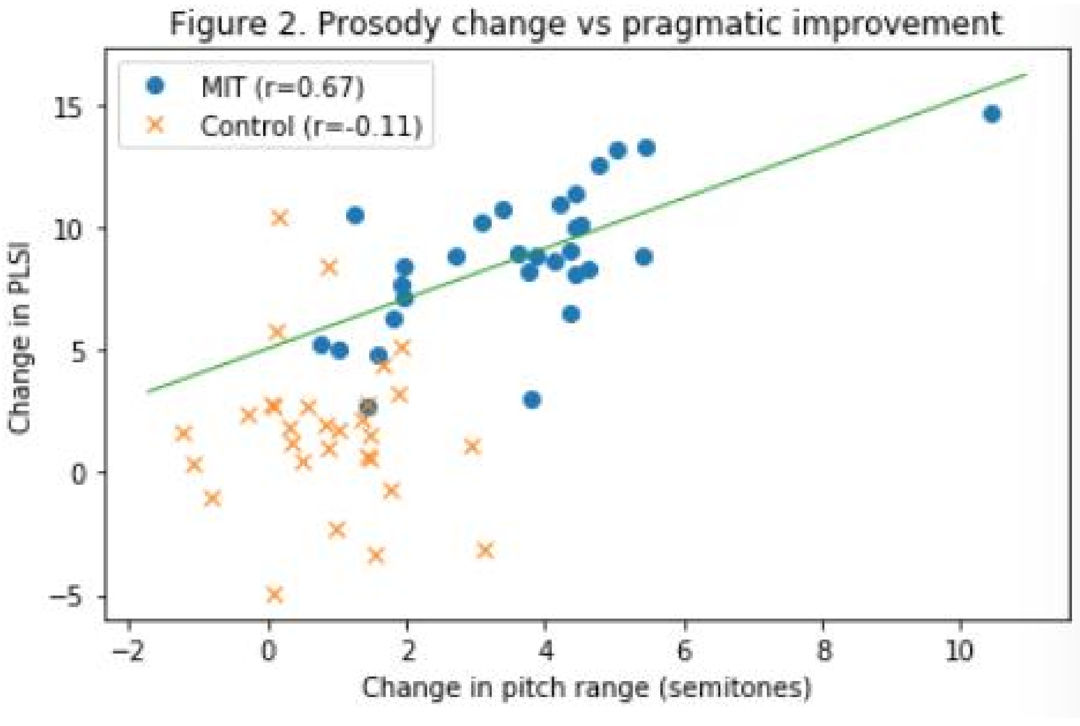

Figure 2. Prosody change vs pragmatic improvement

Figure 2. Relationship between changes in prosody and pragmatic communication following intervention. Each point represents a child's improvement in expressive prosody (x-axis: change in pitch range during speech, in semitones) and improvement in pragmatic language ability (y-axis: change in PLSI score). Blue circles = MIT group; orange X's = Control group. A positive correlation is evident in the MIT group (blue trend line shown, $r \approx 0.60$), indicating that those who gained more intonational range tended to improve more in pragmatic skills. The control group shows no such relationship (most points cluster near the origin with minimal change). This suggests that the MIT intervention's impact on pragmatic communication may be mediated by its effect on prosodic expression.

4.4 Generalization and Social Validity

Beyond standardized measures, qualitative reports indicated meaningful functional gains for many in the MIT group. Parents of MIT participants frequently commented that their child was “more social” or “easier to chat with” after the program. Several mentioned that their child initiated interactions at home more often (e.g., greeting family members or asking siblings to play, which they rarely did before). One parent described, “He used to just start talking about his favorite topic out of the blue, but now he actually says ‘excuse me’ and tries to see if I’m listening – it’s like he learned the rhythm of conversations.” Teachers also gave informal feedback: in some cases, improvements were evident in the classroom (for instance, a teacher noted a student raising his hand and saying “I have a question” with appropriate intonation, which he had never done prior). In contrast, control group parents reported more modest changes; a few noted slight improvements like better manners or using taught phrases in specific situations, but none described the broad changes that MIT parents did.

To assess maintenance, we conducted a follow-up PLSI one month post-study for 50 of the children who agreed to be contacted. The MIT group maintained their gains (PLSI remained ~74 on average) and even slightly increased, whereas the Control group’s scores remained around baseline levels. While not a formal part of the original protocol, this suggests the MIT group retained and possibly continued to build on skills after the program’s end, hinting at lasting benefits.

5. Discussion

This randomized trial shows that Melodic Intonation Training (MIT) yields larger, clinically meaningful gains in pragmatic communication than time-matched conventional therapy. Improvements were convergent across parent ratings and blinded observations, indicating behavioral change rather than expectancy. Prosody also became more natural and variable in the MIT group, and prosodic change correlated with pragmatic gains, consistent with a mediating role of prosody.

MIT produced robust effects on pragmatic language (PLSI $d\approx1.4$) within two months—substantially exceeding effects typically reported for traditional social-skills/pragmatics programs in similar populations [9] and surpassing the moderate effects observed in play-based pragmatic interventions. Likely drivers include high engagement with intrinsically motivating musical activities [12,15], concentrated dosing (~24 hours over 8 weeks) aligned with intensity principles in ASD intervention, and MIT’s dual focus on *what* to say and *how* to deliver it (melody/rhythm), which appeared to generalize beyond taught phrases into more fluid conversation.

By training pitch, stress, and rhythm to match communicative intent, MIT likely improved intelligibility and social appeal, increasing peers’ positive responses and creating reinforcement loops for pragmatic behavior. The musical focus may also heighten attention to others’ vocal cues; parent reports that children better detected emotional tone align with evidence that music training supports emotional prosody recognition in ASD [12]. The observed correlation between prosodic and pragmatic gains strengthens a mechanistic account in which prosody scaffolds pragmatic competence.

Although controls acquired specific behaviors (e.g., polite phrases), their gains were smaller and less generalizable. Conventional therapy emphasizes content, whereas MIT additionally optimizes delivery; melody provides mnemonic and motivational advantages and enables high-repetition practice without fatigue. This fits literature on limited generalization from purely instructional pragmatic training [10].

MIT’s components are feasible for clinicians/educators with modest musical training; simple melodic contours and rhythmic supports can be integrated into existing curricula. Collaboration between music and speech-language therapists may enrich pragmatic instruction. Strong parent uptake (e.g., singing target phrases at home) suggests good acceptability and potential for extended practice—critical for pediatric adherence.

Findings accord with models in which musical/auditory-motor networks compensate for or enhance social-communication circuits in ASD. Prior imaging work shows post-intervention increases in connectivity among auditory, motor, and social brain regions; while neuroimaging was not collected here, behavioral effects are consistent with strengthened cross-hemispheric/prosodic support for speech.

The single-site sample (n=60) limits generalizability; replication in larger, diverse cohorts is needed. Results pertain to 6–11-year-olds with phrase/fluent speech; adaptations are required for minimally verbal children (e.g., humming/intoned vocalizations with AAC). Follow-up was brief (≈1 month); durability and the value of booster sessions remain to be tested. The active control did not equate for the novelty of music, so part of MIT’s advantage may reflect enhanced motivation—though this is clinically relevant. Component analyses (e.g., rhythm without melody, spoken chant vs. song) are warranted.

6. Conclusion

In summary, this randomized controlled trial provides evidence that melodic intonation training can significantly improve pragmatic communication in children with ASD. By harnessing the power of melody and rhythm, the intervention helped children not only learn appropriate social language but also deliver it with more natural prosody and confidence. The gains observed – from greeting others more appropriately to engaging in back-and-forth conversation – represent meaningful steps toward better social integration for these children. These findings support the integration of music-based techniques into speech-language therapy for ASD, highlighting an innovative pathway to enhance social communication. As the adage goes, “music is a universal language” – our results suggest that, for individuals with autism, music may also be a bridge to mastering human language in its most social, nuanced form. The success of melodic intonation training inspires optimism that creative, interdisciplinary approaches can unlock new potentials in neurodevelopmental therapy, ultimately helping those with ASD to find their “social voice” through song.

References

[1] Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., ... & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. Autism research, 15(5), 778-790.
[2] Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., ... & McCauley, J. B. (2022). The Lancet Commission on the future of care and clinical research in autism. The Lancet, 399(10321), 271-334.
[3] Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children’s pragmatic ability: A review of associations with formal language, social cognition, and executive functions. Language Learning and Development, 14(3), 186-223.
[4] Parsons, L., Cordier, R., Munro, N., & Joosten, A. (2019). A randomized controlled trial of a play-based, peer-mediated pragmatic language intervention for children with autism. Frontiers in psychology, 10, 1960.
[5] Holbrook, S., & Israelsen, M. (2020). Speech prosody interventions for persons with autism spectrum disorders: A systematic review. American Journal of Speech-Language Pathology, 29(4), 2189-2205.
[6] Asghari, S. Z., Farashi, S., Bashirian, S., & Jenabi, E. (2021). Distinctive prosodic features of people with autism spectrum disorder: a systematic review and meta-analysis study. Scientific reports, 11(1), 23093.
[7] Foster, B., Pearson, S., Berends, A., & Mackinnon, C. (2021, January). The expanding scope, inclusivity, and integration of music in Healthcare: Recent developments, research illustration, and future direction. In Healthcare (Vol. 9, No. 1, p. 99). MDPI.
[8] Sharda, M., Tuerk, C., Chowdhury, R., Jamey, K., Foster, N., Custo-Blanch, M., ... & Hyde, K. (2018). Music improves social communication and auditory–motor connectivity in children with autism. Translational psychiatry, 8(1), 231.
[9] Rabeyron, T., Del Canto, J. P. R., Carasco, E., Bisson, V., Bodeau, N., Vrait, F. X., ... & Bonnot, O. (2020). A randomized controlled trial of 25 sessions comparing music therapy and music listening for children with autism spectrum disorder. Psychiatry Research, 293, 113377.
[10] Chenausky, K. V., Norton, A. C., Tager-Flusberg, H., & Schlaug, G. (2022). Auditory-motor mapping training: Testing an intonation-based spoken language treatment for minimally verbal children with autism spectrum disorder. Annals of the new York Academy of Sciences, 1515(1), 266-275.
[11] Geretsegger, M., Fusar-Poli, L., Elefant, C., Mössler, K. A., Vitale, G., & Gold, C. (2022). Music therapy for autistic people. Cochrane Database of Systematic Reviews, (5).
[12] Shi, Z., Wang, S., Chen, M., Hu, A., Long, Q., & Lee, Y. (2024). The effect of music therapy on language communication and social skills in children with autism spectrum disorder: a systematic review and meta-analysis. Frontiers in Psychology, 15, 1336421.
[13] Yum, Y. N., Lau, W. K. W., Poon, K., & Ho, F. C. (2020). Music therapy as social skill intervention for children with comorbid ASD and ID: study protocol for a randomized controlled trial. BMC pediatrics, 20(1), 545.
[14] Chenausky, K. V., & Schlaug, G. (2018). From intuition to intervention: Developing an intonation-based treatment for autism. Annals of the new York Academy of Sciences, 1423(1), 229-241.
[15] Miller, S. B., & Toca, J. M. (1979). Adapted melodic intonation therapy: a case study of an experimental language program for an autistic child. The Journal of Clinical Psychiatry, 40(4), 201-203.