Article

CHILD LANGUAGE TEACHING & THERAPY

Child Language Teaching and Therapy
2022, Vol. 38(2) 180–196

DOI: 10.1177/02656590221080308
journals.sagepub.com/home/clt
SAGE

Effect of Sing and Speak 4 Kids: An Online Music-Based Speech and Language Learning Game for Children in Early Intervention

Hayoung A Lim
Department of Music, Oral Roberts University, Tulsa, Oklahoma, USA

Erica M Ellis
Department of Communication Disorders, California State University, Los Angeles, California, USA

David Sonnenschein
iQsonics LLC, Santa Monica, CA, USA

Abstract

Music-based speech language interventions have shown promise to support young children with autism, other speech and language deficits, and Dual Language Learners (also known as DLL, English Language Learners, or ELL). Online edtech learning programs may produce greater positive outcomes for children by including parents as mediators of the intervention. This study measured the preliminary effectiveness of Sing and Speak 4 Kids (SS4Kids), a music-based online speech and language development game, administered to 26 children ages 2–6 years old with or at risk for a diagnosis of autism, other speech and language deficits or DLL. The children were trained in early intervention settings across 4–6 sessions over a 2-week period in one of three group conditions: (a) teacher only in clinic; (b) parent only at home; (c) both teacher + parent. Measurement of verbal production of target words in pre- and post-training sessions showed that trained words significantly improved from pre-test to post-test Additionally, there was no effect of different group conditions (teacher only vs. parent only vs. both teacher + parent) on children's performance. Results suggest that the SS4Kids program is an effective music-based speech and language training method for supporting target word production in young children across a two week timespan. Importantly, the results also found that group conditions did not influence the improvement, confirming effectiveness of both clinic and home-based parent mediation. During a time when traditional in-person intervention services may be restricted, the current work provides cautious but emerging evidence of the effectiveness of an online edtech evidence-based practice to support the speech and language outcomes for a variety of children in early intervention.

Corresponding author:
David Sonnenschein, iQsonics LLC, 1127 18th St #4, Santa Monica, CA 90403, USA.
Email: david@iqsonics.com

Keywords
speech, language, music therapy, intervention, telepractice, online learning, edtech, autism, children, parents

Introduction

Due to the Coronavirus pandemic that began in 2020, there has been an increased and accelerated use of digital online service delivery, also known as telepractice, in the education and healthcare fields, including those that support for young children with developmental disabilities, disorders, or at risk for developmental delays and their families. Personnel in educational systems, early childhood health care providers, clinical practitioners and families have all had to adjust to new or alternative methods to support the needs of their students and children (Hao et al., 2021; Tambyraja et al., 2021). More than 1 million U.S. preschool children diagnosed with Autism Spectrum Disorder (ASD), Intellectual Disability (ID), or "at risk" are impacted by speech-language delays (CDC, 2020). Children with speech and language delays or disorders often have difficulty learning and using words, following directions, answering questions, and participating in conversations depending on the level of severity of their diagnosis. The usage of digital technology can make pediatric speech therapy more engaging and provide multiple opportunities for multisensory learning (Jesus et al., 2019).

Speech & language learning

Children with or at risk for developmental disorders are often in need of early intervention services due to delays in their communication abilities (Centers for Disease Control and Prevention, 2021). Weaknesses can be seen in prelinguistic as well as linguistic domains. For example, young children who are typically developing achieve several communication milestones within the first two years of their lives, while those who have delays may have difficulty with early communication skills such as joint attention, babbling, and later demonstrate difficulty in word learning abilities (Oller, 2014). Some children are able to learn early language skills, but are limited in their variety or use of words (McGregor et al., 2013). Other children may have more communication skills, but lack in their abilities to follow simple directions, produce fewer complex sounds and sound combinations, and have difficulty using more advanced language (ASHA, n.d.).

In order to learn language (whether vocal as with speech, or manual as with sign language), one must be able to perceive, produce and use symbols or words to understand and communicate. This capacity involves the processing of patterns and the learning of regularities of grammar and syntax, phonemes, and an extensive vocabulary (Vihman et al., 2009). The statistical patterns of languages allow individuals, including infants and children, to track regularities of sounds and sound combinations in their environments from the input they are receiving (Hay et al., 2011).

Music perception

Music is composed of many separate yet interconnected components such as pitch, melody, rhythm, harmony, form, timbre, and dynamics. These elements are typically arranged in patterns and perceived as "music" (Berger, 2002; Radocy and Boyle, 2003). Similar to how one learns language, these musical patterns are organized in such a way as to bring anticipation of incoming patterns over a temporal order and are necessary for perception (Berger, 2002). As soon as information is perceived as structured and organized musical patterns, which the brain prefers to process rather than random items, the brain begins activation in the higher channels of cognition.

Interestingly, similar perceptual processes of speech patterns are at work when perceiving organized patterns in music (Eysenck, 2001; Lim, 2010; Lipscomb, 1996; Radocy and Boyle, 2003). Essentially, the Gestalt perceptual laws as applied to melodic perception suggest that listeners are likely to perceive musical tones close together in time and auditory space as a melodic unit (i.e. proximity). Trehub &Trainor (1993) indicated that pitch-contour processes in music perception and prosodic processes in speech require a common auditory perceptual strategy. For example, rhythm perception (Clarke, 1987; Radocy and Boyle, 2003; Sloboda, 1985), timbre perception and memory of musical elements such as rhythm pattern are based on the perceptual organization of sequential structure. Thus, these elements are similar to learning speech or to memory of events (Radocy and Boyle, 2003). In music, rhythm provides the temporal ordering of sounds and is perceived as a temporal figure based on the beat patterns. Rhythmic patterns are organized by grouping and perceived as a unit. In addition, rhythmic patterns have constancy and repetitive nature (Berger, 2002). The continuity and perseverative nature of rhythm provides the perceptual elements that ultimately facilitate focused cortical activation, since the brain attends to the repetitive nature of pulse interacting with rhythmic patterns (Berger, 2002; Thaut, 2005). Rhythm is considered one of the most important elements in the learning of spoken language (Berger, 2002). The temporal pattern in language embellishes and causes anticipation of the following pattern (Barrett, 1999; Peters, 1983).

Pitch is a psychological property of tones that is perceived categorically (i.e. C, D#, G, etc). Frequency, the physiological property of tone, produces certain patterns of neural excitation in the higher auditory pathway. Listeners tend to perceive these patterns and organize them into a certain category (Radocy and Boyle, 2003). Perceived patterns are stored and recalled as pitch memory. Organized patterns of different pitches constitute a melody in music, as well as prosody of speech. Perception of melody is based on melodic contour which is the overall shape or particular pattern of pitch movements. Melodic contour strengthened by rhythm is commonly perceived as a "Gestalt" or whole figure. Therefore, melody perception follows the principles of Gestalt perceptual organization (Lipscomb, 1996; Radocy and Boyle, 2003). Melodic pattern or contours, such as wide sweeps and leaps from low to high pitches (e.g. "Twinkle, Twinkle, Little Star"), or scale-wise melodies (e.g. "Mary Had a Little Lamb") create different physiological and psychological states of anticipation in listeners (Berger, 2002; Radocy and Boyle, 2003). Therefore, melodic songs can stimulate a center in the brain that analyzes different sequences of pitch and processes the melodic pattern, such as Broca's area. Melodic contour might facilitate the center to intone the prosody of letters or words in songs, and then eventually to lead to proper speech production (Berger, 2002).

Music perception follows the same principles of general perceptual organization, such as pattern recognition and grouping information into categorical units, thus the mechanism of music perception is similar to the perceptual and cognitive mechanism of speech and language information (Radocy and Boyle, 2003; Thaut, 2005). The similarities between music and speech/language (McMullen and Saffran, 2004), suggest a link between the two domains in terms of their development, perceptual and eventual categorization.

Researchers indicate that general auditory processing mechanisms responsible for pattern analysis are involved in the perception of both speech and music. The brains of young children are quite flexible and show a remarkable ability to reorganize auditory events from speech and music, which suggests that early experience or training has a profound effect on the cortical and perceptual organization of auditory stimuli (McMullen and Saffran, 2004). The research literature regarding music and language supports the idea that their roots are indistinguishable (Patel, 2008, 2011). From the early perception of sounds to the emergence of singing and speech, a close relationship between music and language development is evident. It is therefore during the early stages

of development, when music and speech are highly integrated, the closer link between the two domains should be encouraged (Hafteck, 1997). Due to the similarities, children might learn how to both speak and sing in close tandem, incorporating both in a very structurally cohesive way.

Incorporation of music & language in intervention

If speech and language perception along with music lie on the same perceptual paradigm for young children, it is possible to design therapeutic mechanisms to examine how speech stimuli embedded in musical patterns influence outcomes. Children with ASD may have abnormal auditory-cortical activation that causes a dysfunction of specific temporal regions which specialize in the processing of spoken words and the integration of complex sounds (Boddaert et al., 2004). Nevertheless, the majority of children with ASD show an intact ability to perceive and produce complex sounds, including musical sounds (Schuler, 1995; Tager-Flusberg, 1997). Musical stimuli are commonly organized and perceived by patterns, and this musical pattern perception is observed in children with ASD (Heaton et al., 1999; Orr et al., 1998). The perception of music parallels perception of language, and these processes appear to be intact in children with ASD. Both musical stimuli and language stimuli are perceived as patterns in the sensory channels of children with ASD (Berger, 2002; Thaut, 2005), since pattern perception is also the primary mechanism for speech and language in children with ASD (Prizant, 1983). In particular, language development in children with ASD is heavily influenced by their capacity for pattern perception and production (Prizant, 1983; Prizant, et al., 1997).

As in music, in language the presence of patterns is very evident (Berger, 2002). Every word, when divided into its syllabic rhythm, displays patterns. The functions of patterns in speech/language can be found in musical patterns. Berger (2002) suggests that non-verbal children with ASD tend to be more attentive and motivated to imitate and learn word sounds that are broken down into rhythmically patterned syllables, spoken, clapped, and/or sung. A simple verbal phrase repeated to a rhythmic pattern might sustain the children's attention and interest in the verbal input and increase their anticipation of the following phrase (Lim, 2009, 2010).

The ability to analyze auditory information accurately in a melodic pattern is of vital importance in learning speech prosody. Prosody refers to the variation of tones used when speaking (i.e. intonation or pitch) and vocal stress, which is the relative emphasis given to certain syllables in a word (McCann and Peppe, 2003). The functions of prosody are to provide an indication of the speech affect and speaker's intention. Many children with ASD lack speech affect or prosody (Tager-Flusberg, 1997). They are, however, able to imitate melodic patterns in songs and produce prosody in musical speech (Edgerton, 1994; Lim, 2010; Lim and Draper 2011; Wigram, 2000). Children with ASD might be able to express their emotions through musical melodies, and furthermore to produce prosodic vocal self-expression through melodic patterns. The musical preference and sensitivity of autistic children can be positively transferred to their non-musical behaviors by activating the common perceptual mechanisms (Thaut, 2005). Researchers have also reported that children with ASD showed no deficit in processing affective patterns in musical stimuli, in spite of their affective and interpersonal deficits (Heaton et al., 1999). Children with ASD could recognize emotional expression in music at a simple level, and that they might have intact perception of melody patterns in music, such as a key or melodic mode. Furthermore, they appear to have an intact perceptual ability to connect the musical pattern with a corresponding affective pattern, and that they often have a perceptual preference for music (Heaton et al., 1999).

A relatively small number of researchers have explored the influence of music on speech and language skills in children with language impairments, including children with ASD and other

developmental disorders. Buday (1995) investigated the use of music as a strategy to increase short-term memory for manual signs in children with ASD. A regularly repeated musical pattern improved attention span and recalling rate in the children. Buday (1995) suggested that speech elements paired with a melodic pattern as a cue can be produced more easily in children with ASD. These findings indicate a positive effect of music on sign and speech imitation in children with ASD. Music enables a child with autism to focus more intently upon on-task behaviors by reducing boredom and provides a more enjoyable learning situation for many of the children (Buday, 1995).

Brownell (2002) investigated the effect of musical social stories on the behaviors of children with autism in four experimental case studies. Both the spoken version and the sung version were significantly more effective in addressing the target behavior (i.e. following directions and using a quiet voice) than the control condition. Brownell (2002) reported that the frequency of the negative target behaviors, such as excessive talking about TV, displaying difficulty in following directions and using an intensely loud shouting voice, occurred least often during the presentation of the sung social story. The findings suggest that children with ASD might comprehend a message in a song, and that this comprehension might be enhanced by perception of patterns in the presented music.

Lim (2010) explored how the perception of music impacts the perception and production of speech in young children with autism spectrum disorders (ASD) while examining the effect of "Developmental Speech-Language training through Music" (DSLM). Participants' verbal production was measured before and after music or speech training. Participants in the music training made greater progress than participants in the speech training. Both high and low functioning participants improved their speech production after receiving either music or speech training; however, low functioning participants showed a greater improvement after the music training than the speech training. Children with ASD perceived the important linguistic information embedded in music stimuli that were organized by principles of pattern perception.

Lim and Draper. (2011) compared a common form of Applied Behavior Analysis Verbal Behavior (ABA VB) approach and music incorporated with ABA VB method as part of developmental speech-language training in the speech production of children with ASD. This study explored how the perception of musical patterns incorporated in ABA VB operants (mand, tact, echoic, and intra-verbal) impacted the production of speech in children with ASD. Participants included 22 children with ASD between the age of 3 to 5 years who were verbal or pre-verbal with presence of immediate echolalia. Every child was randomly assigned a set of target words for each of the 3 training conditions: (a) music-incorporated ABA VB, (b) speech (ABA VB), and (c) no training. Results showed how both music and speech training were effective for production of the four ABA verbal operants. However, the difference between music and speech training was not statistically different. Music can be incorporated into the ABA VB training method, and musical stimuli can be used as successfully as ABA VB speech training to enhance the functional verbal production in children with ASD (Lim and Draper, 2011).

Sharda et al. (2018) evaluated the neurobehavioural outcomes of a music intervention, compared to a non-music control intervention, on social communication and brain connectivity in 51 school-age children aged 6–12 years with ASD during 8–12 weeks of music or non-music intervention. The music intervention involved use of improvisational approaches through song and rhythm to target social communication. The results indicated that 8–12 weeks of music intervention (relative to non-music behavioral intervention) can improve parent-reported social communication and functional brain connectivity in the children, thus supporting the use of music as a therapeutic tool for individuals with ASD (Sharda et al., 2018).

Music-based speech-language therapy, compared to traditional intervention, is an important avenue to explore to improve outcomes for students with ASD and other speech or language

disorders, based on recent studies linking musical and linguistic skills (Moreno et al., 2011; Strait and Kraus, 2011). For example, children with ASD often show an acute proclivity in their response to music, despite other deficits, and music-based speech language training has been shown effective in improving socio-communicative deficits for children with ASD (Kaplan and Steele, 2005; Kim et al., 2009). Intonation-based treatments such as Auditory-Motor Mapping Training (AMMT), which includes interactive music-making to enhance social interaction and singing to encourage vocal output, offer a promising approach for facilitating expressive language in nonverbal children with ASD (Chenausky et al., 2016; Wan et al., 2010), including generalization (Wan et al., 2011), by stimulating the Mirror Neuron System deficit in ASD speech brain regions through overlap with music-making processing regions. The OPERA hypothesis (Patel, 2011) suggests that musical training supports language learning by enhancing the neural encoding of speech when all these conditions are met:

- **O**verlap – There is anatomical overlap within the brain networks that process an acoustic feature used in both music and speech (e.g. pitch, spectral shape).
- **P**recision – Music training demands greater precision in certain aspects of auditory processing than does ordinary speech perception.
- **E**motion – Engagement with music elicits strong positive emotion and enhances the focus of the learner.
- **R**epetition – Musical training often requires many hours of repeating the same composition.
- **A**ttention – Musical activities that engage this neural network are associated with focused attention.

Considering these studies and theories cited, more work is recommended to explore how treatment and intervention approaches with a music-based foundation can support children's speech and language learning.

Teletherapy in early intervention

Traditionally, early intervention services are provided in person, in an educational, clinic or in-home setting. Many early intervention services are provided with caregivers involved, or in small groups; however, the worldwide pandemic has impacted the options for the services to be provided in the traditional way. While teletherapy (also known as telepractice, or online learning) is not new, there are pros and cons to using teletherapy in early intervention with factors such as access to technology and parent/caregiver support playing a role in a child and family's success and satisfaction (Cole et al., 2019; Fairweather et al., 2016). Options to support learning outcomes include parent coaching (virtual and in-person), home-based programs, as well as computer-based programs and applications (apps). The use of mobile digital devices offers new resources for participation and engagement of children diagnosed with ASD who have an interest in technology and a documented proclivity for the use of technological devices (Gillette, 2003). Innovations are emerging in the efforts to improve outcomes for these socio-communicative deficits, including computerized and gamified interventions (Grynszpan et al., 2014; Zakari et al., 2014), providing extensive player/patient engagement and a channel for delivering measurable health-related changes (Baranowski et al., 2008). Efficacy of telehealth as a service delivery model to coach parents on intervention strategies for their children's early communication skills has been demonstrated (Simicek et al., 2017), as well as for speech-language pathologists working remotely during the COVID-19 era (Beiting and Nicolet, 2020). A literature review (Chen et al., 2016) reported the effectiveness of computers as a more engaging type of speech-language intervention with more tools to enrich the intervention programs, particularly when it comes to children, and significantly

effective in training people with a variety of speech disorders. The success of each of these areas suggests that the combination of telepractice, music therapy and game-based interventions could be even more effective for young children with speech and language impairments.

Sing and Speak 4 Kids

Few studies exist examining the intersection between music therapy and online game-based interventions for children with varying disorders, as well as the impact in different intervention settings. To research this area, iQsonics developed and launched *Sing and Speak 4 Kids* (SS4Kids), a music-based speech and language game, based on Lim's (2010) study of DSLM. Using Lim's protocol, six songs with a total of 36 target vocabulary words are embedded in Ruby on Rails (https://rubyonrails.org/) programming software and played online on any mobile or laptop device (Android, iOS, Windows or MacOS). The songs and target words are based on developmentally appropriate vocabulary and social situations and are programmed into varying levels of game play. Responses (i.e. "Yes/No" answers) by participants are inputted on the touchscreen by the teacher or parent.

The beginning stages of learning each song rely upon the exact repetition of what has been heard (echolalia), a function common to both individuals with ASD and typically developing children learning to speak. First, the onscreen singer is accompanied by musical instrumentation, then in the next stage sings alone. A photo (one of three randomized to encourage generalization) representing the target word at the end of the phrase is presented simultaneously to the singing of the target word. The next stages of the program drop out the last word of each lyric phrase, requesting for the child to verbally fill-in the blank target word, and then repeats the same phrase with speaking only, no melody, and the same request The final stage presents the photo and asks the child to verbalize the correct associated word. See Figure 1 of stages of the SS4Kids Play Protocol.

The training program implements scaffolding and fading of support systems so the child can better attain generalization, spontaneous communication, and independence of the program. The program uses theories and practice of emerging interventions, as well as combines expertise drawn from the fields of communication disorders, music therapy, early childhood development, autism research, educational technology and game-based learning. The program itself is unique in that it specializes in the usage of music therapy as a source of intervention and is optimized for both onsite (clinical, educational) and home-based settings, and includes train-the-trainer materials, analytics, and self-practice tools. While the SS4Kids program has been in development for several years, there is a significant need to examine the effectiveness of evidenced-based, innovative programs that are currently available for early intervention, especially since the onset of a worldwide pandemic.

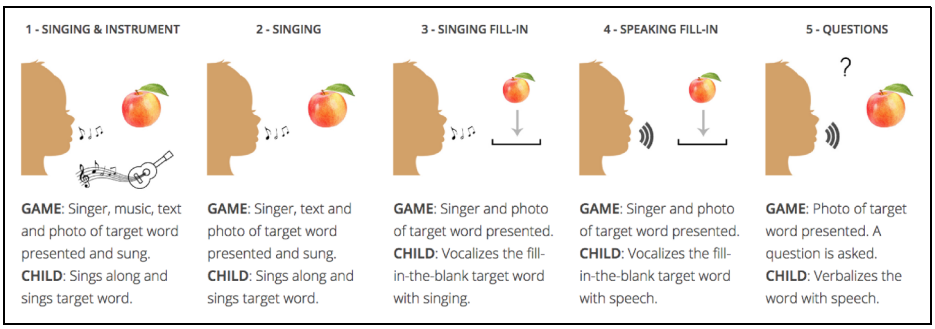


Figure 1. SS4Kids playing protocol.

Purpose

The purpose of the study was to measure the preliminary effectiveness of Sing and Speak 4 Kids in young children in early intervention settings. The study examined the children's verbal production of target words before the training (pre-test), after the last training session (immediate post-test) as well as the effect of different training settings (training with teacher vs. training with parent vs. training with both teacher and parent) on the performance of the children in early intervention settings. We asked the following research questions:

1. Does Training condition (words Trained vs. Untrained) influence the score changes of pre-test to post-test?
2. Does Group condition (Teacher vs. Parent vs. Teacher + Parent) influence the score changes of pre-test to post-test score?
3. Is there any interaction between Training condition and Group condition on the score changes in pre- to post-test scores?

Method

Study design

This study used a randomized, single-blind (outcome assessor), pre- and post-test, within subject design with 3 Group conditions (Teacher in clinic, Parent in home, and Teacher + Parent) for the trained words (a set of 12 target words) and untrained words (a set of 6 target words). Participants were randomly split into 3 groups: Teacher only, Parent only, Teacher + Parent and trained for 4–6 sessions (average 5.3 sessions) of 5–15 min each, over a 2-week period. Outcomes of participants' target word production were measured by comparing pre-test and post-test scores with two independent variables: (1) words trained vs. untrained; and (2) Teacher only vs. Parent only vs. Teacher + Parent. The control condition was the song-word training condition (i.e. 3 songs tested, but only 2 randomized songs trained). The 12 trained words of 2 songs had a maximum score of 48 (if all words were scored as "correct") and the 6 untrained words of 1 song had a maximum score of 24 during the pre- and post-test The dependent variable was the score changes between pre-test and post-test

Participants

A total of 26 children ages 2–6 years old (mean age = 2.88) participated in the study. Children had previous diagnoses of ASD (DSM-5), or identified as "at risk" for impairment. Some children were also identified as Dual Language Learners (DLLs) (indicated by parent report), with or without speech and language disorders. Recruitment was through outreach at local early intervention centers and California State Regional Centers, as well as via flyers and email communication. To meet the inclusion criteria, children were between 2–6 years of age, had access to home internet, were previously identified with ASD or at risk for ASD, other speech and language disorders or delays, and/or speech impairments, and/or learning English as their second language (also known as Dual Language Learners, DLL, English Language Learners, or ELL), and were in an early intervention setting. Children with hearing or vision impairment, and/or recurrent aggressive behavior were not included in the study.

In this study, 14 participants were 2 years old, 6 participants were 3 years old, 3 participants were 4 years old, one participant was 5 years old, and 2 participants were 6 years old. There were 16 boys and 10 girls participating in this study. Sixteen participants had a previous diagnosis of ASD, at risk or parent reported speech and language impairments. Thirteen participants were reported to be Dual Language

Learners (DLLs), including 6 participants that were DLL with impairments based upon parent report. Parents indicated a variety of languages used in their homes, including Arabic, Armenian, Gujarati, Spanish, and Tagalog. See Table 1 for an overview of participant characteristics.

Procedure

Once eligibility was determined, parents completed various intake forms, including a demographic form and the Oral Roberts University IRB approved parent consent form. Parents reported on their child's diagnosis; however, details about current level of communication was minimal. During intake parents answered a simple questionnaire about their child, including a question about current communication abilities. Over half of the parents (N = 16) reported their child had speech and/or language difficulties. Additional speech and vocabulary testing (EVT-3 and GFTA-3) was attempted by a blind and unbiased assessor, when appropriate[1]. However, for many of the participants, the standardized vocabulary testing did not result in valid scores due to language differences and/or being outside the age range.

Testing was completed at one of two early intervention centers in Los Angeles County, while training sessions were completed in the centers and/or at the family home, depending on Condition. Specifically, the Teacher setting was in one of the local early intervention centers, while Parent setting was in the family's home. For Teacher + Parent setting, the first and last sessions were in the center, the 4 middle sessions were at home. Each participant was assessed at two times on his/her ability to produce the 18 target words: (1) Pre- before the first training session, to establish baseline responses; (2) Post- following the last session, to examine the progress. Trained professionals, blind and unbiased to the training conditions for each of the target words, conducted the assessments.

Measures

The full SS4Kids program includes 6 songs with embedded pre- and post- assessments; for the present study, only 3 of the 6 songs were selected, and the pre- and post-assessments were separated from the song play portion of the program so they could be used by blind, unbiased assessors. Each participant was assessed pre- and post-trial on the 18 target words of the 3 songs. The participant was trained in 2 of the 3 randomized songs, with a total of 12 target words for each child. Participants were trained for 4 to 6 (average 5.3) sessions with a duration minimum of 5 min and maximum 15 min, over a 2-week trial period. If participants did not complete at least 4 training sessions, they were not included in the analysis (26 of the 43 enrolled children completed the

Table 1. Summary of participant characteristics.

Age in years	# Participants	Diagnosis/ Impairment	Language status	Impairment + DLL
2	14 (5 F, 9 M)	1 ASD, 7 Other	8 DLL, 6 monolingual-English	4
3	6 (3 F, 3 M)	1 ASD, 1 Other	4 DLL, 2 monolingual-English	1
4	3 (2 F, 1 M)	3 ASD, 0 Other	1 DLL, 2 monolingual-English	1
5	1 (0 F, 1 M)	1 ASD, 0 Other	0 DLL, 1 monolingual-English	0
6	2 (0 F, 2 M)	2 ASD, 0 Other	0 DLL, 2 monolingual-English	0
Average = 2.88	Total = 26 (10 F, 16 M)	8 ASD, 8 Other	13 DLL, 13 monolingual-English	6

Notes. Other = Undiagnosed, at-risk, or parent reported impairment; DLL = Dual Language Learner as reported by parent.

minimum training required for completion of study). Trainers were surveyed about why the participants did not complete at least 4 sessions (i.e. teacher, parent and/or student related reasons, to be considered for increasing usability and fidelity of implementation). The total of 26 participants were randomly allocated into 3 groups: T - Teacher only (n = 8), P - Parent only (n = 7); T + P - Teacher + Parent (n = 11). Outcomes of participants' target word production were measured by comparing pre-test and post-test scores with two independent variables: (1) words trained vs. not trained; and (2) Teacher only vs. Parent only vs. Teacher + Parent.

In all Training and Group conditions, children were assessed on their ability to verbally produce the 18 target words. The assessment tool used was the SS4Kids Likert Scale, an in-program rating system (No Response, Incorrect, Approximation, Correct) which utilized the same image prompt for both pre- and post-assessment (See Figure 2). During this assessment, the Assessor read aloud the question on the screen (e.g. "What is this?"), while the program showed an accompanying photo illustrating the correct target word. The Assessor then rated the child's response in the program interface, giving them two prompts if needed with approximately ten seconds each before proceeding to the next word.

To facilitate generalization of the expressive language outside the intervention session, song lyrics are designed to incorporate objects, actions and situations representative of the child's developmentally appropriate everyday life (e.g. eating, playing, dressing, etc.). Each target word in the program is represented by one of three randomized photos, to increase the generalization effect. The parent is trained to bring the online game experience into the real world by incorporating the lyrics and target words throughout the day and reinforce this activity with other family members and environments such as school.

Statistical analysis

Demographic characteristics taken before the study were compared using Fisher's Exact test to see whether the characteristics differ at the outset. The group differences with respect to the biometric features (Child Types of DLL, Impairment conditions, Autism, Gender, Age, Primary Language) were

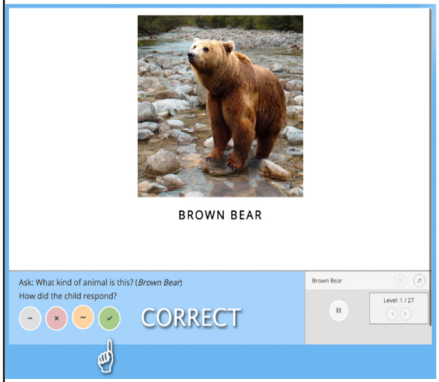
Example	CHILD'S RESPONSE	ASSESSMENT
	Child didn't say anything after being asked the question twice	No Response " - "
	Child answered with a word or sound that was not recognized by the assessor as the target word	Incorrect " X "
	Child spoke a word that was not pronounced 100% correctly, but was similar to the target word (e.g. "app" for "apple")	Approximation " ~ "
	Child properly pronounced the target word	Correct " ✓ "

Figure 2. SS4Kids Likert scale response assessments example and criteria.

compared based on the overall pre-study score using nonparametric Kruskal-Wallis H tests. Second, the study's primary outcome, namely improvement on the post-study score over the study period, was investigated using paired t-tests at the end of the study. The primary outcome was measured by the difference score (i.e. score change in pre-study and post-study scores) for both trained and untrained words. Third, ANOVA and ANCOVA tests were conducted for examining group differences in the score change with respect to other demographic features. In addition, a mixed-effect linear regression model was used to investigate if there were any significant interactions among demographic features in explaining the difference score and the post-study score on both trained and untrained words.

Results

Results of the pre-test indicate that there were individual differences in the speech and language skills before participating in the SS4Kids program, based on participant's age, gender, diagnosis of autism, speech/language impairments, and DLL. The post-test scores indicate these differences could be counteracted while engaging in the SS4Kids program, therefore benefitting all participants. The in-program Likert Scale Rating system (No Response; Incorrect; Approximation; Correct) utilized a previously exposed image during post-assessment and had inter-rater reliability of 93.25% and validity established at $p < 0.05$. This result confirms the reliability and validity ($p < 0.05$) of the pre-and post-test in the present study to measure the level of speech production in children who used the SS4Kids program.

Nineteen of the 26 participants (73%) had an improvement in overall post-study scores. In terms of trained and untrained groups, 17 out of 26 participants (65%) had an increase in post-study scores with trained words, while only 14 out of 26 participants (54%) had an increase in post-study score with untrained words. Paired t-tests were conducted to evaluate the score difference between pre- and post-test stages in terms of the overall score, trained score, and untrained score (See Table 2). The average overall post-test score ($M = 50.54$, $SD = 16.25$) was significantly higher than the pre-test score ($M = 44.7$, $SD = 14.89$). The results indicated a significant effect of the SS4Kids program overall and on trained words, therefore confirming the primary outcome of the program. The post-test score ($M = 34.23$, $SD = 11.31$, $p = 0.025$) for trained words significantly improved compared to the pre-test score ($M = 30.27$, $SD = 10.41$). However, there was no significant difference between pre-test score ($M = 14.5$, $SD = 5.34$) and post-test scores ($M = 16.31$, $SD = 5.58$, $p = 0.062$) for untrained words.

Another t-test was conducted on the difference in the average of post-test scores with trained ($M = 34.23$, $SD = 11.31$) and untrained words ($M = 16.31$, $SD = 5.58$) to examine the effect of training in the study. The test results ($t = 7.81$, $p = 0.00$) showed that training had a significant effect on the test score compared to untrained words. Additionally, the effect size appeared to be moderate between overall pre-test and post-test scores ($d = 0.655$), where the main contribution came from the SS4Kids program with trained words ($d = 0.609$). However, the effect size between untrained scores was classified as small ($d = 0.468$). This result indicates that participants who received the

Table 2. Analysis of SS4Kids program effect for trained and untrained words; T-test of difference.

	Study Arms	Mean difference	*SE*	*t*	p
Program Effect Post-test vs. Pre-test	Trained and Untrained words	5.808*	2.20	2.592	0.016
	Trained words	3.96*	1.63	2.384	0.025
	Untrained words	1.808	0.91	1.95	0.06
Training Design	Trained vs. untrained words	0.814*	0.10	7.81	0.00

Note. N = 26. * The mean difference is significant (two-tailed tests) at the 0.05 level.

training via SS4Kids improved their verbal production of the target words. This particular result implies a positive effect of a two-week training via SS4Kids on the verbal production of target words in young children (2–6 year old) with speech/language impairments and/or DLL.

The analysis showed that the group conditions (Teacher vs. Parent vs. Teacher + Parent) did not influence the pre-test score ($M = 29.94$, $p = 0.123$) and post-test score ($M = 34.00$, $p = .155$). The Teacher + Parent group had higher scores ($M = 36.91$) than the other two groups ($M = 28.38$ (Teacher), and $M = 36.71$ (Parent)); however, the difference was not statistically significant. In addition, the results from the ANOVA showed no significant relationship among the training type (Trained vs. Untrained), group conditions (Teacher vs. Parent vs. Teacher + Parent), and mean difference (score change) between pre- and post-test This means that SS4Kids can be effective and beneficial for children in any of the listed conditions and/or settings. Teacher in clinic, parent in home, and teacher with parent can result in the improvement of target word production using the SS4Kids program.

The parameter estimates obtained from the mixed-effect linear regression model showed that Child Type (with three levels: Impairment, DLL & Impairment, and DLL Only), condition of impairment, and availability of official diagnosis were significant factors in influencing the study outcomes. However, the interactions between each two of the biometric features were not significant. Based on the analysis, results obtained from multiple ANOVA tests, the improvements on overall post-test scores were significant in participants with impairment and DLL conditions ($F = 15.65$; $p = 0.001$) and participants identified as DLL Only ($F = 5.76$; $p = 0.033$), while the score change is not significant for participants with only impairment. The results indicate that participants in the group of DLL only had higher scores of post-test than those with speech or language impairment or DLL and impairment $F\ (2, 20) = 5.76$, $p = .033$; *Partial Eta Square* = .312, and that was statistically significant. This result might indicate the feasibility of using SS4Kids for supporting general speech and language skills for children who are DLLs without any other speech and language concern or developmental diagnosis.

Discussion

Overall, the results of the present study answered the research questions and showed that participants using the SS4Kids program improved their scores on the target word production from pre-test to post-test This finding suggests that a two-week long training via SS4Kids is effective for increasing both vocabulary acquisition (semantics) and speech production in children with speech/language impairments due to ASD and/or other developmental disorders, as well as for children who are DLLs with and without impairments. Even more, the results suggest that the program supports target word production for children with and without disorders, and regardless of whether English is their first language or not.

Importantly, the results from the study found that Group conditions did not influence the pre-test score and post-test score. Specifically, teachers alone in clinic, parents alone in home, and/or teachers with parents as trainers, each show improvement of children's production outcomes. Given the emerging need of educators and families to find new or alternative supports for their students and children, these findings are encouraging.

Since traditional interventions often take place in a classroom or clinic setting, many clinicians and educators are looking for alternative methods and tools to support the learning of their clients and students during the restrictions of the pandemic and anticipated hybrid clinic + home intervention settings in the future. The preliminary work here demonstrated that for children enrolled in early intervention, parents can also support positive outcomes in their child's development by utilizing the program. Early intervention has long emphasized the importance of family and caregiver input and participation. Studies have shown that parents can successfully be taught to use evidence-

based strategies to enhance their child's social-communicative functioning (Oono et al., 2013), parents can integrate with teacher training with positive results (Nieuwboer et al., 2013), and parents, students, and clinicians report satisfaction with telepractice (McCullough, 2001). Additional benefits of parent-mediated intervention include increase in generalization and maintenance of children's language skills (Koegel et al., 1982). The current work suggests that parents, in addition to professionals, can be active facilitators of positive outcomes for their children.

Although very promising, this study possesses a number of limitations that need to be acknowledged and could be used to inform future research. The SS4Kids program seems to be an effective training method for improving target word production, however results should be interpreted with some caution. Limitations to be considered include the study design as a within-subject design, which was chosen to collect as much data as possible from a small group size with all participants being exposed to the intervention (with 2 out of 3 possible songs trained), rather than a between-subject design that would have required control group of participants to receive no training at all. Another limitation pertains to the participant samples in terms of both their sizes and their composition; more homogeneous participant profiles (e.g. autism only or DLL only) might result in more distinct outcomes for each type of population, leading to more distinct and preferable recommendations and training conditions for each population. Further, the quantitative measurement of outcome was restricted to the verbal production of the intervention target words; it is recommended that future research include (a) expanded measurements of generalization of verbal production beyond the target words to a greater vocabulary usage, (b) maintenance of verbal production gains in delayed post training periods of a month or more, and (c) measurement of other social communication behaviors besides verbal production (e.g. eye contact, facial expressions, gestures, integration of vocal/non-vocal, social overtures, social responses, etc.) that might be improved.

While the results of the study demonstrated an increase in the verbal production of target words from pre- to post-test, only parent reported details of pre-intervention communication skills of participants were collected due to challenges with obtaining standardized vocabulary measures, as well as little information about participant language exposure (for DLLs). Future work in the area should obtain detailed information about language exposure and linguistic history by using a tool such as the LEAT (DeAnda et al., 2016) as well as measure participant communication skills using alternative and dynamic methods of assessment prior to the intervention. Additionally, the study had a relatively small sample size and a short time span of the study, which limits the generalizability of results. Therefore, these conclusions must be interpreted with some caution.

Conclusion

During a time when traditional in-person intervention services may be restricted, the current work provides emerging evidence of an effective tool to support the speech and language outcomes for a variety of children in early intervention. The goal of SS4Kids is to provide an option for families and practitioners that is accessible and fun. The program uses theoretical evidence along with a novel strategy that offers a combination of music, gameplay, digital delivery to any internet connected device and low cost While the evidenced-based effectiveness of SS4Kids is demonstrated in the current work, the research is ongoing as more research needs to be examined. In the future, feedback from parents, practitioners and children should be explored to examine conditions for program usability, feasibility and fidelity of implementation, as well as effectiveness in maximizing performance outcomes. In the meantime, the program can be useful as a tool to support existing strategies during remote service delivery for families and provides promise for enhancing the traditional treatment approaches for children impacted by speech and language impairments in early intervention.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Xxxxxxx.Dr. Hayoung Lim and Dr. Erica Ellis received consulting fees from the IES / Dept. of Education SBIR funding for the research study presented in this article. David Sonnenschein is the CEO of iQsonics LLC, developer of the edtech program Sing and Speak 4 Kids, the intervention studied in this article, which was funded by IES / Dept. of Education SBIR for this research study.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Institute of Education Sciences, (grant number 91990019C0023).

ORCID iDs

Erica M. Ellis https://orcid.org/0000-0001-7904-4660
David Sonnenschein https://orcid.org/0000-0003-2416-8126

Note

1. Standardized testing was not conducted with children outside of the age range and/or learning more than one language, since it would not be an accurate assessment of their speech and language skills. While limited, the mean standard scores (SS) and ranges for those participants able to complete testing are EVT-5 (N = 3), SS M = 78, SS range = 72- 84; GFTA-3 (N=9), SS M= 89, SS range = 50–111.

References

American Speech Language Hearing Association (ASHA). (n.d.). Early identification of speech, language, and hearing disorders. Retrieved January 25, 2021, from https://www.asha.org/public/early-identification-of-speech-language-and-hearing-disorders/#identify

Baranowski T, Buday R, Thompson DI, and Baranowski J (2008) Playing for Real: Video Games and Stories for Health-Related Behavior Change. *American Journal of Preventive Medicine*, 2008 January, 34(1): 74–82.

Barrett M (1999) *The Development of Language*. UK: Psychology Press Ltd

Beiting M and Nicolet G (2020) Screenless teletherapy and silent telesupervision: leveraging technology for innovative service delivery and clinician training in speech-language pathology during the COVID-19 Era. *CommonHealth* 1(3): 106–120

Berger DS (2002) *Music Therapy, Sensory Integration and the Autistic Child*. Jessica Kingsley Publishers Ltd: UK.

Boddaert N, Chabane N, Belin P, et al. (2004) Perception of complex sounds in autism: Abnormal auditory cortical processing in children. *American Journal of Psychiatry* 161: 2117–2120.

Brownell MD (2002) Musically adapted social stories to modify behaviors in students with autism: Four case studies. *Journal of Music Therapy* 39: 117–144.

Buday EM (1995) The effects of signed and spoken words taught with music on sign and speech imitation by children with autism. *Journal of Music Therapy* 32: 189–202.

Centers for Disease Control and Prevention (2020) Data & statistics on autism spectrumCenters for disease control and prevention (2020). Data & statistics on autism spectrum disorder, Page 1, https://www.cdc.gov/ncbddd/autism/data.html

Centers for Disease Control and Prevention (2021, April) National Center on Birth Defects and Developmental Disabilities: Why Act Early if You're Concerned about Development? https://ectacenter.org/~pdfs/pubs/importanceofearlyintervention.pdf

Chen YP, Johnson C, Lalbakhsh P, et al. (2016) Systematic review of virtual speech therapists for speech disorders. *Computer Speech & Language* 37: 98–128, ISSN 0885-2308,

Chenausky K, Norton A, Tager-Flusberg H, et al. (2016) Auditory-Motor mapping training: comparing the effects of a novel speech treatment to a control treatment for minimally verbal children with autism. *PLoS ONE* 11(11): e0164930

Clarke EF (1987) Categorical rhythm perception: An ecological perspective. In. Gabrielsson A (Ed.) *Action and Perception in Rhythm in Music*. Stockholm: Royal Swedish Academy of Music, 19–33.

Cole B, Pickard K, and Stredler-Brown A (2019) Report on the Use of telehealth in early intervention in Colorado: strengths and challenges with telehealth as a service delivery method. *International Journal of Telerehabilitation* 11(1): 33–40

DeAnda S, Bosch L, Poulin-Dubois D, et al. (2016) The language exposure assessment tool: quantifying language exposure in infants and children. *Journal of Speech, Language, and Hearing Research* 59(6): 1346–1356

Edgerton CL (1994) The effect of improvisational music therapy on the communicative behaviors of autistic children. *Journal of Music Therapy* 31: 31–62.

Eysenck MW (2001). *Principles of Cognitive Psychology*. (2nd ed.). Philadelphia, PA: Psychology Press Ltd

Fairweather CG, Lincoln MA, and Ramsden R (2016) Speech-language pathology teletherapy in rural and remote educational settings: decreasing service inequities. *International Journal of Speech-Language Pathology* 18(6): 592–602..

Gillette D (2003) Technology's role in encouraging social and cognitive development in children with autism. *Journal of Developmental and Learning Disabilities* 35: 97–109.

Grynszpan O, Weiss PL, Perez-Diaz F, et al. (2014) Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism* 18(4): 346–361.

Hafteck L (1997) Music and language development in early childhood: integrating past research in the two domains. *Early Child Developmental and Care* 130: 85–97.

Hao Y, Franco JH, Sundarrajan M, et al. (2021) A pilot study comparing tele-therapy and in-person therapy: perspectives from parent-mediated intervention for children with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 51(1): 129–143

Hay JF, Pelucchi B, Graf Estes K, et al. (2011) Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology* 63(2): 93–106

Heaton P, Hermelin B, and Pring L (1999) Can children with autistic spectrum disorder perceive affect in music? An experimental investigation. *Psychological Medicine* 29: 1405–1410.

Jesus LMT, Martinez J, Santos J, et al. (2019) Comparing traditional and tablet-based intervention for children with speech sound disorders: A randomized controlled trial. *American Speech-Language-Hearing Association* 62(11): 4045–4061, https://pubs.asha.org/doi/full/10.1044/2019_JSLHR-S-18-0301

Kaplan RS and Steele AL (2005) An analysis of music therapy program goals and outcomes for clients with diagnoses on the autism spectrum. *Journal of Music Therapy* 42: 2–19.

Kim J, Wigram T, and Gold C (2009) Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. *Autism* 13(4): 389–409.

Koegel RL, Schreibman L, Britten KR, et al. (1982) Comparison of parent-training to direct child treatment. In: Koegel RL, Rincover A, and Egel AL (eds) *Educating and Understanding Autistic Children*. San Diego, CA: College-Hill Press, 260–279.

Lim HA (2009) Use of music to improve speech production in children with autism Spectrum disorders: theoretical orientation. *Music Therapy Perspectives* 27(2): 103–114.

Lim HA (2010) Effect of "developmental speech and language training through music" on speech production in children with autism Spectrum disorders. *Journal of Music Therapy* 47(1): 3–25.

Lim HA and Draper E (2011) The effects of music therapy incorporated with applied. Behavior analysis verbal behavior approach for children with autism Spectrum disorders. *Journal of Music Therapy* 48(4): 532–550.

Lipscomb SD (1996) The cognitive organization of musical sound. In Hodges DA (Ed.) *Handbook of Music Psychology* (2nd ed.). San Antonio. TX: IMR Press, 133–175.

McCann J and Peppe S (2003) Prosody in autism spectrum disorders: A critical review. *International Journal of Language and Communication Disorders* 38(4): 325–350.

McCullough A (2001) Viability and effectiveness of teletherapy for pre-school children with special needs. *International Journal of Language and Communication Disorders* 36(Suppl. 1): 321 – 3326.

McGregor K, Oleson J, Bahnsen A, and Duff D (2013) Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *Int J Lang Commun Disord* 48(3): 307–319. https://doi.org/10.1111/1460-6984.12008. Epub 2013 Feb 25.

McMullen E and Saffran JR (2004) Music and language: A developmental comparison. *Music Perception* 21: 289–311.

Moreno S, Bialystok E, Barac R, et al. (2011) Short-term music training enhances verbal intelligence and executive function. *Psychological Science* 22(11): 1425–1433.

Nieuwboer CC, Fukkink RG, and Hermanns JMA (2013) Online programs as tools to improve parenting: A meta-analytic review. *Children and Youth Services Review* 35(11): 1823–1829.

Oller JWJr (2014) Milestones in language development. In Brooks PJ and Kempe V (eds.) *Encyclopedia of Language Development*. Thousand Oaks, CA: SAGE Publications, Inc., 377–381. https://dx.doi.org/10.4135/9781483346441.n123

Oono IP, Honey EJ, and McConachie H (2013) Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *The Cochrane Database of Systematic Reviews* 30;(4): CD009774.

Orr TJ, Myles BS, and Carlson JK (1998) The impact of rhythmic entertainment on a person with autism. *Focus on Autism and Other Developmental Disabilities* 13: 163–166.

Patel A (2008) *Music, Language, and the Brain*. Oxford: Oxford University Press.

Patel A (2011) Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol* 2: 142. 29 June 2011 https://doi.org/10.3389/fpsyg.2011.00142

Peters AM (1983) *The Units of Language Acquisition*. London: Cambridge University Press.

Prizant BM (1983) Language acquisition and communicative behavior in autism: toward an understanding of the "whole" of it. *Journal of Speech and Hearing Disorders* 48: 296–307.

Prizant BM, Schuler AL, Wetherby AM, et al. (1997) Enhancing language and communication development: language approaches. In Cohen D and Volkmar F (Eds.) *Handbook of Autism and Pervasive Developmental Disorders*. New York: John Wiley &Sons, Inc, 572–605.

Radocy RE and Boyle JD (2003) *Psychological Foundation of Musical Behavior* (4th ed.). Springfield, IL: Charles C Thomas Publisher. Ltd

Schuler AL (1995) Thinking in autism: Differences in learning and development. In Quill K (Ed.) *Teaching children with Autism: Methods to Enhance Communication and Socialization*. Albany, NY: Delmar, 11–32.

Simacek J, Dimian AF, and McComas JJ (2017). Communication intervention for young children with severe neurodevelopmental disabilities Via telehealth. *Journal Of Autism And Developmental Disorders* 47(3): 744–767

Sharda M, Tuerk C, Chowdhury R, et al. (2018) Music improves social communication and auditory–motor connectivity in children with autism. *Translational Psychiatry* 2018; 8:231.

Sloboda JA (1985) *The Musical Mind*. Oxford, UK: Clarendon Press.

Strait D and Kraus N (2011) Can You hear Me Now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology* 2011(2): 113.

Tager-Flusberg H (1997) Perspectives on language and communication in autism. In Cohen D and Volkmar F (Eds.) *Handbook of Autism and Pervasive Developmental Disorders*. New York: John Wiley &Sons, Inc, 894 – 8900.

Tambyraja SR, Farquharson K, and Coleman J (2021) Speech-Language teletherapy services for school-aged children in the United States during the COVID-19 pandemic. *Journal of Education for Students Placed at Risk (JESPAR)* 26(2): Special issue on Responses to educating students at risk during the COVID-19 Pandemic; Guest Editor: Sinéad Harmey, Oxfordshire: A publication of Taylor & Francis Group, LLC, 91–111. Published online: 01 Apr 2021. https://doi.org/10.1080/10824669.2021.1906249

Thaut MH (2005) *Rhythm, Music, and the Brain: Scientific Foundations and Clinical Applications*. New York: Routledge.

Trehub SE and Trainor LJ (1993) Listening strategies in infancy: The roots of music and language development. In McAdams S and Bigand E (Eds.) *Thinking in Sound: The Cognitive Psychology of Hearing and Audition*. Oxford: OUP, 278–327.

Vihman MM, DePaolis RA, and Keren-Portnoy T (2009) Babbling and words: A dynamic systems perspective on phonological development. In: Bavin EL (Ed.) *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press, 163–182.

Wan CY, Demaine K, Zipse L, et al. (2010) From music making to speaking: Engaging the mirror neuron system in autism. *Brain Research Bulletin* 82(3–4): 161–168.

Wan CY, Bazen L, Baars R, et al. (2011). Auditory-Motor mapping training as an intervention to facilitate speech output in Non-verbal children with autism: A proof of concept study. *PLoS ONE* 6(9): e25505

Wigram T (2000) A method of music therapy assessment for the diagnosis of autism and communication disorders in children. *Music Therapy Perspectives* 18: 13–18.

Zakari HM, Ma M, and Simmons D (2014) A review of serious games for children with autism Spectrum disorders (ASD). In: Ma M, Oliveira MF, and Baalsrud Hauge J (eds) *Serious Games Development and Applications. SGDA 2014. Lecture Notes in Computer Science*. 93–106. Cham: Springer.