



# Auditory-Motor Mapping Training in a More Verbal Child with Autism

Karen V. Chenausky<sup>1,2</sup>, Andrea C. Norton<sup>1</sup> and Gottfried Schlaug<sup>1,2\*</sup>

<sup>1</sup>Music, Neuroimaging, and Stroke Recovery Laboratory, Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA, United States, <sup>2</sup>Department of Neurology, Harvard Medical School, Boston, MA, United States

We tested the effect of Auditory-Motor Mapping Training (AMMT), a novel, intonation-based treatment for spoken language originally developed for minimally verbal (MV) children with autism, on a more-verbal child with autism. We compared this child's performance after 25 therapy sessions with that of: (1) a child matched on age, autism severity, and expressive language level who received 25 sessions of a non-intonation-based control treatment Speech Repetition Therapy (SRT); and (2) a matched pair of MV children (one of whom received AMMT; the other, SRT). We found a significant Time  $\times$  Treatment effect in favor of AMMT for number of Syllables Correct and Consonants Correct per stimulus for both pairs of children, as well as a significant Time  $\times$  Treatment effect in favor of AMMT for number of Vowels Correct per stimulus for the more-verbal pair. Magnitudes of the difference in post-treatment performance between AMMT and SRT, adjusted for Baseline differences, were: (a) larger for the more-verbal pair than for the MV pair; and (b) associated with very large effect sizes (Cohen's  $d > 1.3$ ) in the more-verbal pair. Results hold promise for the efficacy of AMMT for improving spoken language production in more-verbal children with autism as well as their MV peers and suggest hypotheses about brain function that are testable in both correlational and causal behavioral-imaging studies.

## OPEN ACCESS

### Edited by:

Ana Inês Ansaldo,  
Université de Montréal, Canada

### Reviewed by:

Noman Naseer,  
Air University, Pakistan  
Diana López-Barroso,  
University of Málaga, Spain  
Ignacio Moreno-Torres,  
University of Málaga, Spain

### \*Correspondence:

Gottfried Schlaug  
gschlaug@bidmc.harvard.edu

**Received:** 18 April 2017

**Accepted:** 09 August 2017

**Published:** 04 September 2017

### Citation:

Chenausky KV, Norton AC and Schlaug G (2017) Auditory-Motor Mapping Training in a More Verbal Child with Autism.  
*Front. Hum. Neurosci.* 11:426.  
doi: 10.3389/fnhum.2017.00426

**Keywords:** autism, speech therapy, intonation, AMMT, minimally verbal, speech development

## INTRODUCTION

Autism spectrum disorder (ASD), a neurodevelopmental condition affecting approximately 1/68 children (Christensen et al., 2016), is characterized by: (1) impairment in social communication; and (2) the presence of repetitive behaviors or restricted interests (American Psychiatric Association and American Psychiatric Association. Task Force on DSM-5, 2013). While many children with ASD have language within the normal range (Kjelgaard and Tager-Flusberg, 2001; Tager-Flusberg et al., 2005), Rapin et al. (2009) identified two subgroups of children with ASD and impaired expressive language: one whose receptive language is also impaired, and the other whose receptive language is intact. In addition, 25%–46% of children who receive a diagnosis of ASD remain minimally verbal (MV) past age five (Tager-Flusberg et al., 2005; Kasari et al., 2013; Tager-Flusberg and Kasari, 2013; Rose et al., 2016) meaning that they have an expressive vocabulary smaller than 20 words and no word combinations (Kasari et al., 2013). Thus, spoken language treatment is important for many children with ASD.

In contrast to language skills, which are severely disordered in minimally-verbal children with ASD, musical ability may be relatively intact. For example, Applebaum et al. (1979)

compared the ability of three autistic teens with no musical training to that of three typically-developing teens with considerable musical experience (e.g., at least 4 years of piano lessons) on a musical perception/production task. Although participants were matched in age, the average IQ of the autistic participants was 68, more than two standard deviations below average. Each participant was asked to imitate tone sequences of increasing complexity, ranging from single pitches to sets of four tones in an atonal configuration. The autistic participants performed as well as or better than the typically-developing, musically-trained participants on 62% of trials on average (range 45%–90%). In other work, Lai et al. (2012) found no differences in parent ratings of musical affinity in MV autistic children compared to typical controls, despite the autistic participants' severely disordered language. Music has also been used therapeutically with MV children and adults with ASD. For example, Boso et al. (2007) reported that 52 weekly, hour-long music therapy sessions were associated with both significantly improved music skills (e.g., singing melodies of differing lengths) and decreased clinical severity scores. Taken together, the relative strengths of severely-affected individuals with ASD in the areas of musical interest, ability and learning suggest that musical activities may be a productive medium through which to develop and foster communication skills for this population.

We recently introduced and demonstrated the efficacy of a novel intonation-based treatment for MV children with ASD, Auditory-Motor Mapping Training (AMMT), which has been used successfully in that population (Wan et al., 2011) and shown to outperform a non-intonation-based control treatment, Speech Repetition Therapy (SRT; Chenausky et al., 2016, 2017). AMMT is a modification of Melodic Intonation Therapy (MIT), which has been used successfully to improve speech production in left-hemisphere stroke patients with severe nonfluent aphasia (Schlaug et al., 2009, 2010; Zipse et al., 2012). In this case report, we discuss the effects of AMMT on two children with ASD, one MV and one more verbal, and compare them to matched participants receiving the control treatment. Our goal is to understand whether AMMT can produce improvements in spoken language in a more-verbal child, commensurate with those seen in MV children with ASD. We discuss the results in the context of previous imaging findings relating spoken language performance and integrity of two white-matter tracts involved in language processing.

## MATERIALS AND METHODS

### Participants

Four children with ASD, all male, aged between 4 years 1 month and 6 years 7 months, participated in this study. Diagnosis was confirmed by an Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999) score greater than 12. Inclusion criteria were the ability to: (1) correctly repeat at least two speech sounds; (2) participate in table-top activities for at least 15 min at a time; (3) follow one-step commands; and (4) imitate simple gross- and oral-motor movements like clapping hands and opening mouth. Exclusion criteria included a diagnosis of a major sensorineural

or developmental condition other than ASD (e.g., concomitant deafness; known genetic disorders). MV status for the two MV participants was defined as using fewer than 20 intelligible words and having no productive syntax; this was confirmed both by parent report and child performance during initial assessments. The MV participants had scores of 1 (AMMT participant) and 3 (SRT participant) for “words used and understood” on the MacArthur-Bates Communication Development Inventory (MCDI; Fenson et al., 1993) at Baseline. Neither child produced word combinations during assessment sessions, which included obtaining language samples at Baseline. By contrast, the two more-verbal children had scores of 90 (AMMT participant) and 131 (SRT participant) on the MCDI at Baseline. Both more-verbal participants produced word combinations during a language sample at Baseline: the more-verbal AMMT participant, for example, said “I want squeeze it”; the more-verbal SRT participant said “I want more”, “I want” and “I more”. According to the vocabulary and grammar benchmarks laid out in Tager-Flusberg et al. (2009), the two MV participants met criteria for the “First Words” stage (defined as at least 2–15 words and no word combinations) and the two more-verbal participants met criteria for the “Word Combinations” stage (at least 10–50 words and mean length of utterance in morphemes 1.1–2.4).

Two tests determined the number of speech sounds children were able to repeat at baseline: (1) the first two sections of the Kaufman Speech Praxis Test (KSPT; Kaufman, 1995); or (2) a phonetic inventory test where children were asked to imitate 21 consonants and 10 vowels of English. All participants were tested with the Visual Reception, Fine Motor, Receptive Language and Expressive Language subtests of the Mullen Scales of Early Learning (MSEL; Mullen, 1995). MSEL subtest scores are generally reported as *T*-scores ( $\mu = 50$ ,  $\sigma = 10$ ); however, in **Table 1** we report raw scores, which are more informative for this population. All raw scores in **Table 1** correspond to *T*-scores  $< 20$ , classified as “very low”. Thus, though we describe them as “more verbal”, the two non-MV participants in this study still experienced significant language delays. However, they differed from the MV participants on vocabulary size, expressive language score and developmental language stage. The MV and more-verbal AMMT participants were matched to an MV and a more-verbal control participant, respectively, based on a combination of age, ADOS score, expressive language score and phonetic inventory.

Children were recruited from multiple autism centers serving Greater Boston. The study was carried out in accordance with the Declaration of Helsinki and with the approval of the Institutional Review Board of Beth Israel Deaconess Medical Center. Parents of all participants provided written informed consent for their child's participation prior to enrollment. While in the study, children continued with their regular school programs but did not participate in any other speech therapy activities or new treatments outside of school.

### Treatment

In this study, we compared the recently-introduced intonation-based treatment for MV children with ASD, AMMT, which involves repetition of intoned (sung) words or phrases

**TABLE 1** | Participant characteristics.

	Age (year;month)	ADOS <sup>1</sup>	MSEL RL <sup>2</sup>	MSEL EL <sup>3</sup>	MSEL VR <sup>4</sup>	MSEL FM <sup>5</sup>	Phonetic inventory <sup>6</sup>	KSPT §1 <sup>7</sup>	KSPT §2
AMMT-V	4;1	20	13	16	26	22	25	11	63
Control 1: SRT-V	5;3	19	29	18	46	39	23	11	29
Control 2: AMMT-MV	4;2	19	10	9	13	19	11	9	10
Control 3: SRT-MV	6;7	20	12	10	22	21	4	8	3

AMMT, Auditory-Motor Mapping Training; SRT, Speech Repetition Therapy. "V" refers to the more-verbal participants and "MV" to the minimally verbal participants. <sup>1</sup>ADOS: Autism Diagnostic Observation Schedule (Lord et al., 1999) score >12 to confirm ASD diagnosis. <sup>2</sup>RL: Receptive Language subscale of the Mullen Early Learning Scales (MSEL). <sup>3</sup>EL: Expressive Language subscale of the MSEL. <sup>4</sup>VR: Visual Reception subscale of the MSEL. <sup>5</sup>FM: Fine Motor subscale of the MSEL. <sup>6</sup>Phonetic inventory: the number of English vowels and consonants a child was able to imitate (max = 31 phonemes). <sup>7</sup>KSPT: Kaufman Speech Praxis Test (Kaufman, 1995). Max score on Section 1 = 11; max score on Section 2 = 63.

(Wan et al., 2011), to a non-intonation-based control treatment, SRT. The two pitches on which AMMT stimuli are intoned follow a simplified prosodic contour (stressed syllables are intoned on the higher pitch, unstressed syllables on the lower pitch), and all syllables are produced at a rate of approximately one per second. As each syllable is produced, therapist and child tap electronic drums tuned to the same two pitches. AMMT's multimodal nature facilitates spoken language production by activating shared motor, auditory and visual neural representations of the same vocal/manual actions (Meister et al., 2003; Ozdemir et al., 2006; Lahav et al., 2007), mimicking the co-occurrence of babbling and bimanual banging seen in typical development (Iverson and Fagan, 2004; Gernsbacher et al., 2008).

While SRT uses the same stimuli and has the same basic structure as AMMT (see "Treatment Session Structure" Section, below), in SRT stimuli are spoken, not sung; and there is no tapping on drums. Previously, we showed AMMT to produce superior improvements in spoken language in MV children with ASD over SRT (Chenausky et al., 2016, 2017). Here, we investigate whether AMMT can also lead to improvements in a more-verbal child with ASD, compared to SRT.

## Stimuli

Stimuli for the MV children consisted of two sets (Trained and Untrained) of 15 familiar bisyllabic words or phrases referring to people ("mommy"), social actions ("bye-bye") and objects ("bubbles") common in children's daily lives. For the more-verbal children, stimuli also consisted of two sets (Trained and Untrained) of 16 bi- or trisyllabic stimuli, again relevant to children's daily activities (e.g., "bathroom", "computer", "go outside"). Trained stimuli were presented during baseline and probe assessments and practiced during therapy sessions. Untrained stimuli were presented only during baseline and probe assessments to ascertain how well children's speech production skills generalized to unpracticed words and phrases.

## Treatment Session Structure

Treatment sessions lasted approximately 45 min/day, 5 days/week. Breaks and rewards were provided after every 5 to 10 items, as needed. In both treatments, stimuli were presented in the context of the five-step prompt hierarchy outlined below:

1. *Listening*: Therapist introduces target phrase by showing a picture and using the phrase in a semantically meaningful context: "It's fun to blow *bubbles*". Therapist produces target.
2. *Unison*: "Let's do it together: *bubbles*". Therapist produces target with child.
3. *Unison fade*: "Again, together: *bu...*". Therapist produces initial portion of target with child; child continues independently.
4. *Imitation*: (a) "My turn: *bubbles*". Therapist produces target alone. (b) "Your turn:..." Child produces target independently.
5. *Cloze*: "Last time: It's fun to blow..." Therapist presents semantic context for stimulus; child fills in the blank independently.

## Outcome Measures and Assessments

We used four outcome measures in this study. Our primary measure was Syllables Approximated per Stimulus, a global indicator of developing speech. For a syllable to be considered *approximated* (approximately correct), its initial consonant must share two of three phonetic features (manner of articulation, place of articulation and voicing) with the target consonant; and its vowel must share two features (backness, height) with the target vowel. "Backness" and "height" refer to the anterior/posterior and dorsal/ventral position of the tongue in the mouth during production of the vowel. A more detailed explanation of this rubric appears in Chenausky et al. (2016). Three additional measures were also used: Syllables Correct per Stimulus, Consonants Correct per Stimulus and Vowels Correct per Stimulus. A consonant or vowel was considered correct if it matched the target; a syllable was considered correct if its initial consonant and its vowel both matched the target. These more stringent measures assessed speech production precision and accuracy.

Note that the measures in this study differed slightly from those in previous work. In Chenausky et al. (2016), we calculated (e.g.) the percentage of Syllables Approximated out of a total of 60 syllables in both sets of stimuli, then compared mean scores across treatment groups. Here, we compared one child to another. Because our statistical analyses required comparisons of samples with a mean and variance, we modified the measure to Syllables Approximated per Stimulus (yielding a within-subject mean and variance) and compared means between children.

## Transcription Reliability

All baseline and probe responses were phonetically transcribed and scored by coders blind to timepoint. Five Baseline sessions were administered to each child in order to establish a stable level of performance and to allow them time to acclimatize to the room and investigator. A child's Best Baseline was defined as the pre-treatment assessment session during which he produced the most approximately-correct syllables. For no participant was the Best Baseline session the last one, so there was no improvement over the five Baseline sessions. The Best Baseline score was then compared to performance at the P25 (post 25 treatments) assessment. A reliability study, in which 10% of Baseline and probe sessions were independently transcribed by two investigators, yielded a Cohen's  $\kappa = 0.547$ ,  $p < 0.0005$  and 70.1% agreement on consonants correct, and Cohen's  $\kappa = 0.270$ ,  $p < 0.0005$  and 54.7% agreement for vowels correct, as reported in Chenausky et al. (2016). These figures compare favorably to those for infant babbles (Davis and MacNeilage, 1995).

## Statistical Analyses

Two-way repeated-measures analysis of variances (ANOVAs) were performed on each outcome measure with Time (Baseline vs. P25) and Stimulus Type (Trained vs. Untrained) as within-subjects factors, and Treatment (AMMT vs. SRT) as a between-subjects factor. Because some scores were significantly different between paired children at Baseline, total Baseline score was included as a covariate, where appropriate, to correct for differences in performance. Table 2 shows Baseline scores for all participants.

## RESULTS

### More-Verbal Participants

#### Syllables Approximated per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time on Syllables Approximated per Stimulus,  $F_{(1,29)} = 5.677$ ,  $p = 0.024$ . Mean number of Syllables Approximated per Stimulus, averaged over both of the more-verbal participants and both stimulus types and adjusted for Baseline performance, increased over 25 sessions. There were no other significant main effects or two- or three-way interactions for this measure; however, the Time  $\times$  Treatment interaction approached

significance,  $F_{(1,29)} = 3.251$ ,  $p = 0.082$ . The more-verbal AMMT participant improved by an adjusted mean of 0.3 Syllables Approximated per Stimulus, while the more-verbal SRT participant decreased by an adjusted mean of 0.1.

#### Syllables Correct per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time on Syllables Correct per Stimulus,  $F_{(1,29)} = 15.704$ ,  $p < 0.0005$ . Mean number of Syllables Correct per Stimulus, averaged over both more-verbal participants and both stimulus types and adjusted for Baseline performance, increased over 25 sessions. There were no other significant main effects for this measure.

There was a significant Time  $\times$  Treatment interaction for Syllables Correct per Stimulus, adjusted for Baseline performance,  $F_{(1,29)} = 27.787$ ,  $p < 0.0005$ , Cohen's  $d = 1.92$  (very large). The more-verbal AMMT participant improved by an adjusted mean of 0.7 Syllables Correct per Stimulus, while the more-verbal SRT participant decreased by an adjusted mean of 0.1. There were no other significant 2-way interactions; however, there was a significant Time  $\times$  Stimulus  $\times$  Treatment interaction,  $F_{(1,29)} = 5.052$ ,  $p = 0.032$ . The more-verbal AMMT participant improved by an adjusted mean of 0.5 Syllables Correct per Trained Stimulus and by 0.9 Syllables Correct per Untrained Stimulus. The more-verbal SRT participant improved by an adjusted mean of 0.1 Syllables Correct per Trained Stimulus and decreased by 0.3 Syllables Correct per Untrained Stimulus.

#### Consonants Correct per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time on Consonants Correct per Stimulus,  $F_{(1,29)} = 13.505$ ,  $p = 0.001$ . Mean number of Consonants Correct per Stimulus, averaged over both more-verbal participants and both stimulus types and adjusted for Baseline performance, increased over 25 sessions. There were no other significant main effects for this measure.

There was a significant Time  $\times$  Treatment interaction for Consonants Correct per Stimulus, adjusted for Baseline performance,  $F_{(1,29)} = 18.203$ ,  $p < 0.0005$ , Cohen's  $d = 1.56$  (very large). The more-verbal AMMT participant improved by an adjusted mean of 1.2 Consonants Correct per Stimulus, while the more-verbal SRT participant decreased by an adjusted mean of 0.4. There were no other significant two- or three-way effects for this measure.

TABLE 2 | Baseline performance.

	Syllables approximated per stimulus	Syllables correct per stimulus	Consonants correct per stimulus	Vowels correct per stimulus
AMMT-V	1.97	1.25	2.75	1.56
Control 1: SRT-V	0.97 ( $p < 0.0005^1$ )	0.5 ( $p < 0.0005$ )	0.81 ( $p < 0.0005$ )	1.16 ( $p = 0.027$ )
Control 2: AMMT-MV	0.77	0.17	0.63	0.60
Control 3: SRT-MV	0.27 ( $p = 0.003$ )	0.10 (n.s.)	0.23 ( $p = 0.02$ )	0.23 ( $p = 0.015$ )

AMMT, Auditory-Motor Mapping Training; SRT, Speech Repetition Therapy. "V" refers to the more-verbal participants and "MV" to the minimally verbal participants.  
<sup>1</sup> $p$ -values vs. respective AMMT participant.

## Vowels Correct per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time on Vowels Correct per Stimulus,  $F_{(1,29)} = 21.657$ ,  $p < 0.0005$ . Mean number of Vowels Correct per Stimulus, averaged over both more-verbal participants and both stimulus types and adjusted for Baseline performance, increased over 25 sessions. There were no other significant main effects for this measure.

There was a significant Time  $\times$  Treatment effect on Vowels Correct per Stimulus, adjusted for Baseline performance,  $F_{(1,29)} = 13.663$ ,  $p = 0.001$ , Cohen's  $d = 1.35$  (very large). The more-verbal AMMT participant improved by an adjusted mean of 0.5 Vowels Correct per Stimulus, while the more-verbal SRT participant decreased by an adjusted mean of 0.3, over 25 sessions. There were no other significant two- or three-way interactions for this measure.

**Figure 1** shows the change over time on each measure for the two more-verbal participants, adjusted for Baseline performance.

## Minimally Verbal (MV) Participants

### Syllables Approximated per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time for Syllables Approximated per Stimulus,  $F_{(1,27)} = 4.069$ ,  $p = 0.041$ , indicating that the mean number of Syllables Approximated per Stimulus, averaged over both MV participants and both stimulus types and adjusted for Baseline performance, increased over 25 sessions. There were no other significant main effects for Syllables Approximated per Stimulus.

There was a significant Time  $\times$  Treatment interaction for Syllables Approximated per Stimulus, adjusted for Baseline performance,  $F_{(1,27)} = 5.362$ ,  $p = 0.028$ , Cohen's  $d = 0.88$  (large). The MV AMMT participant improved by an adjusted mean of 0.6 Syllables Approximated per Stimulus from Best Baseline to P25, and the MV SRT participant improved by an adjusted mean of only 0.1 over the same time period. There were no other significant two- or three-way effects for Syllables Approximated per Stimulus.

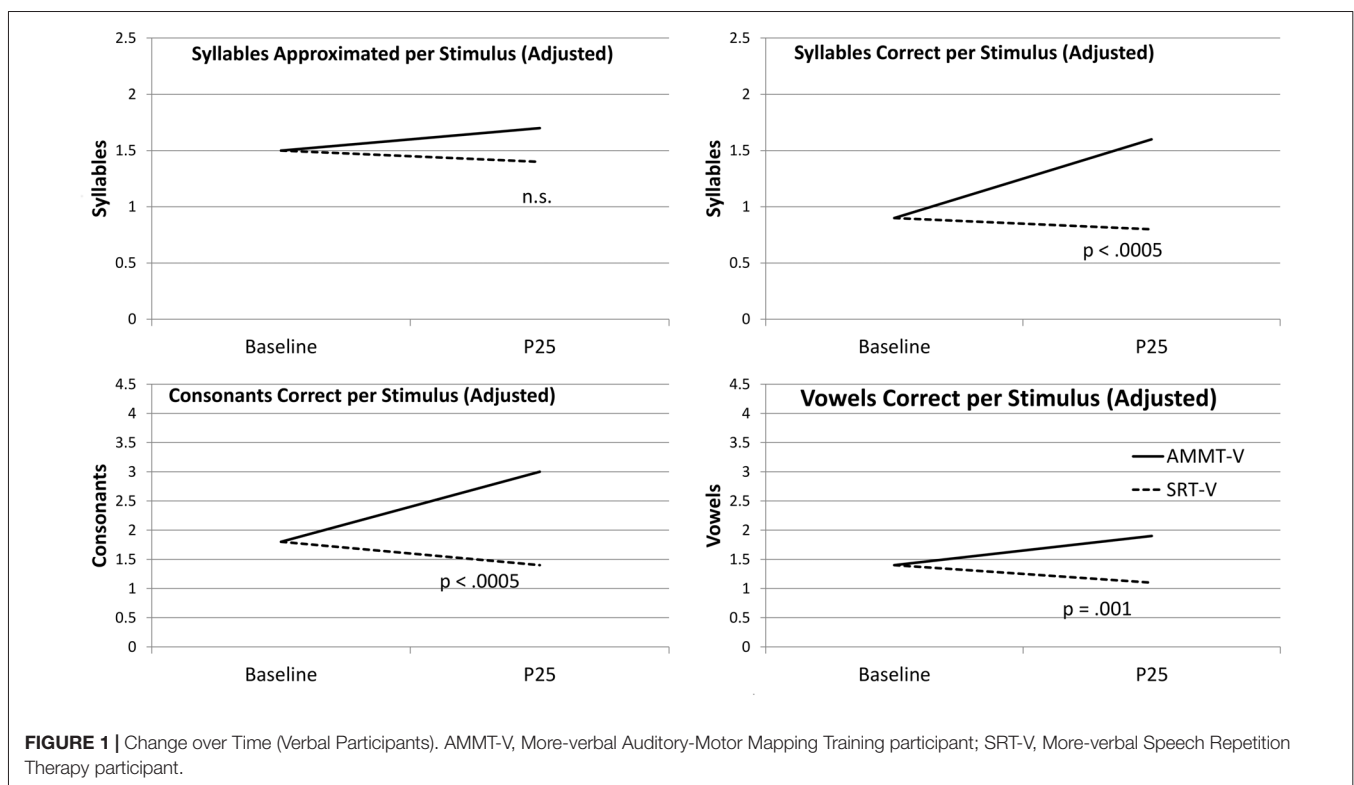
### Syllables Correct per Stimulus

Unadjusted for Baseline performance, there was a significant main effect of Time on Syllables Correct per Stimulus,  $F_{(1,28)} = 16.141$ ,  $p < 0.0005$ . Mean number of Syllables Correct per Stimulus, averaged over both MV participants and both stimulus types, increased over 25 sessions. There were no other significant main effects for this measure.

There was also a significant Time  $\times$  Treatment interaction for Syllables Correct per Stimulus, unadjusted for Baseline performance,  $F_{(1,28)} = 5.271$ ,  $p = 0.029$ , Cohen's  $d = 0.87$  (large). The MV AMMT participant improved by a mean of 0.3 Syllables Correct per Stimulus, and the MV SRT participant by a mean of 0.1 Syllables Correct per Stimulus, over 25 sessions.

### Consonants Correct per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time for Consonants Correct per Stimulus,  $F_{(1,27)} = 8.056$ ,  $p = 0.009$ . Mean number of Consonants Correct per Stimulus, averaged over both MV participants and both stimulus types and adjusted



for Baseline performance, increased over 25 sessions. There were no other significant main effects for this measure.

There was also a significant Time  $\times$  Treatment interaction for Consonants Correct per Stimulus, adjusted for Baseline performance,  $F_{(1,27)} = 5.726$ ,  $p = 0.024$ , Cohen's  $d = 0.9$  (large). The MV AMMT participant improved by an adjusted mean of 0.7 Consonants Correct per Stimulus over 25 sessions, compared to an adjusted mean of 0.1 for the MV SRT participant. There were no other significant two- or three-way interactions for this measure.

### Vowels Correct per Stimulus

Adjusted for Baseline performance, there was a significant main effect of Time for Vowels Correct per Stimulus,  $F_{(1,27)} = 4.506$ ,  $p = 0.043$ . In this case, both MV children improved on this measure over 25 sessions, the MV AMMT participant by 0.5 Vowels per Stimulus and the MV SRT participant by 0.4 (between-treatment difference n.s.). There were no other significant main effects, and no significant two- or three-way interactions, for this measure.

**Figure 2** shows the change over time on each measure for the two MV participants, adjusted for Baseline performance where appropriate.

## DISCUSSION

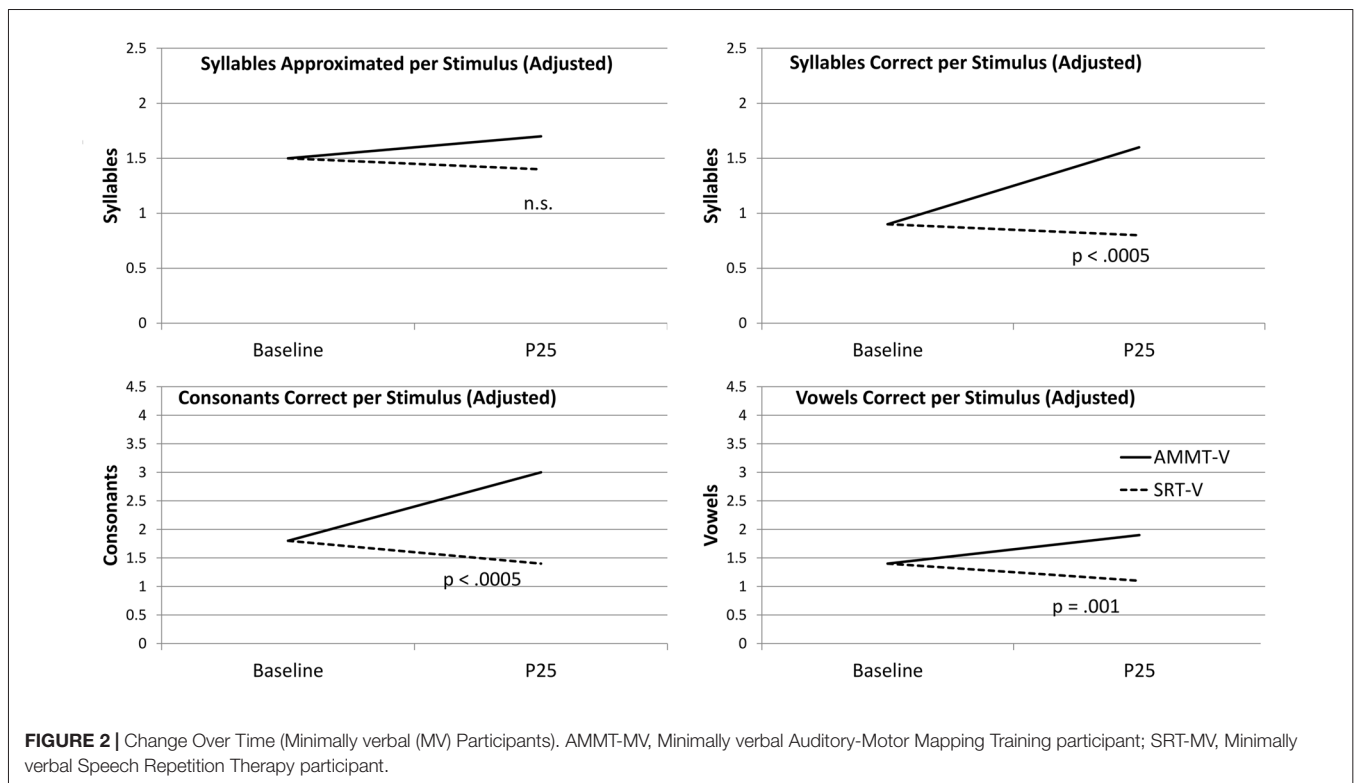
In this study, we assessed the effect of AMMT, an intonation-based spoken language treatment with proven efficacy as an intervention for MV children with ASD (Wan et al., 2011;

Chenausky et al., 2016), on a more-verbal child with ASD. We compared his progress to a matched participant who received SRT and to a matched pair of MV children (one who received AMMT; the other, SRT) with ASD. Several results emerged.

First, AMMT resulted in a greater improvement in most of our outcome measures than SRT for our more-verbal participant, when differences in Baseline performance were taken into account. The more-verbal AMMT participant also experienced a comparatively greater improvement than the more-verbal SRT participant in number of Syllables, Consonants and Vowels Correct per Stimulus, over 25 treatment sessions.

Second, the adjusted improvements seen in AMMT over SRT for Syllables Correct, Consonants Correct and Vowels Correct per Stimulus for the more-verbal participant with ASD are associated with very large effect sizes (Cohen's  $d > 1.3$ ). These effect sizes were larger than those for the difference between the MV AMMT and SRT participants (Cohen's  $d$  approximately 0.9). Together, these results suggest that AMMT may be at least as effective for more-verbal children with ASD as it is for MV children with ASD.

The finding of no significant between-treatment difference in Syllables Approximated per Stimulus suggests that this measure is too coarse-grained for the speech of children with better-than-minimal spoken language. The more stringent outcome measure of Syllables Correct may be more useful for this group of children. Finally, though the differential findings on Trained and Untrained Stimuli for the more-verbal AMMT and SRT participants may be child-specific, it is also possible they may have arisen from differences in phonetic complexity or familiarity of the words/phrases between the two stimulus sets.



Alternatively, they may suggest that AMMT promotes greater skill generalization than SRT. This would be an important finding for children with ASD, who struggle with generalizing skills to new contexts (Plaisted, 2001; Happé and Frith, 2006).

In addition to describing the behavioral results of 25 treatment sessions of AMMT, it is also worth considering the neural substrates likely to be involved in the improvements seen. AMMT is hypothesized to work by engaging a network of brain regions that are activated by auditory, motor and visual actions (Wan et al., 2011). Lai et al. (2012) showed that, while listening to speech, activation in the left inferior frontal gyrus (IFG) was lower in children with ASD compared to typically developing controls, yet it was higher than in controls while listening to song, even in the presence of reduced integrity of the left arcuate fasciculus (AF) in the children with ASD. This suggests that music-making activities and intonation may be a unique vehicle for engaging the IFG in two of its presumed functions—the mapping of sounds to articulatory actions and their sequential execution. The AF connects auditory-perceptual regions in the temporal lobe to motor-related regions in the posterior inferior portion of the frontal lobe (Catani et al., 2005). As such, it is thought to be responsible for the bidirectional mapping of speech articulation and acoustics (Saur et al., 2008; Leclercq et al., 2010), to mediate new word learning (López-Barroso et al., 2013), and to form part of the dorsal language pathway (Friederici, 2011). Further research suggests that integrity of the AF and the extreme capsule fiber tract (EmC), another white-matter tract involved in language comprehension and production, may be related to AMMT participants' improvements in spoken language production. The extreme capsule fiber tract (EmC) links the more anterior portion of the IFG to the middle-posterior portion of the superior temporal gyrus (Makris and Pandya, 2009), assuming a slightly more ventral course for this tract near the insular cortex. Although the functions of the EmC are less clear, it has been proposed that this tract is involved in the comprehension (Saur et al., 2008) and production of morphologically more complex words (Rolheiser et al., 2011).

In previous work (Chenausky et al., 2017) we showed that integrity of the left AF, measured at baseline, predicted improvement in percent syllable-initial consonants correct after 25 sessions of AMMT in a group of 10 MV children with ASD. In ongoing work in our lab, we have found that when FA of the language tracts was relativized by that of the corticospinal tract (CST)—a motor execution network that has been shown to be compromised in ASD (Carper et al., 2015)—relative integrity of the right EmC and both the right and left AF at baseline predicted improvement in percent syllables approximately correct, and relative integrity of the right EmC at baseline predicted improvement in percent vowels correct.

Taken together, the behavioral results reported here, combined with previously identified links between imaging and behavioral findings, lead to empirically testable hypotheses regarding the neural substrates required for more-verbal children with ASD to benefit from an intonation-based treatment. Specifically, we would predict that, in this population

as in the MV group, the degree of improvement in consonant production should be positively related to integrity of the left AF, improvement in vowel production to integrity of the right EmC, and the degree of improvement in syllables correct should be positively related to the integrity of both tracts. Furthermore, the behavioral findings reported on here may support the hypothesis, put forth in Hardy and LaGasse (2013), that external rhythmic, musical cues may provide useful templates for organizing motor output in children with ASD by: (1) decreasing motor planning demands; and (2) increasing movement efficiency and accuracy. Additional support for this hypothesis, as well as for the idea that music is beneficial for children with ASD for reasons beyond its ability to increase motivation and attention, comes from research showing that MIT, the treatment from which AMMT was derived, has been shown to be effective for improving verbal output in patients with moderate to severe nonfluent aphasia after a left-hemisphere stroke (Schlaug et al., 2009, 2010; Zipse et al., 2012; Wan et al., 2014).

## CONCLUDING REMARKS

AMMT holds promise for improving the spoken language of children with ASD who are not considered MV, but who still struggle with significant expressive language and speech-production deficits. The current results should be replicated in a larger group of more-verbal children with ASD. In addition, pre- and post-treatment imaging studies should be performed to both test whether integrity of the left AF and the right EmC are positive predictors of the degree to which more-verbal children with ASD can improve in spoken language production, and to verify whether integrity of those two tracts does, in fact, increase as children's spoken language abilities improve.

## AUTHOR CONTRIBUTIONS

GS, KVC and ACN: substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work. KVC, GS and ACN: drafting the work or revising it critically for important intellectual content. KVC, GS and ACN: final approval of the version to be published. GS, KVC and ACN: agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## FUNDING

Funding for this study was provided by the Nancy Lurie Marks Family Foundation, Autism Speaks and by NIH P50-DC 13027.

## ACKNOWLEDGMENTS

We are deeply grateful for the effort of the dedicated families who committed significant amounts of their time to making this work

possible. We also thank our former Speech-Language Pathology clinical fellows, postdoctoral fellows and junior investigators,

music therapists and all research assistants and summer students for their hard work in the collection of these data.

## REFERENCES

- American Psychiatric Association and American Psychiatric Association. Task Force on DSM-5. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association.
- Applebaum, E., Ege, A., Koegel, R., and Imhoff, B. (1979). Measuring musical abilities of autistic children. *J. Autism Dev. Disord.* 9, 279–285. doi: 10.1007/bf01531742
- Boso, M., Emanuele, E., Minazzi, V., Abbamonte, M., and Politi, P. (2007). Effect of long-term interactive music therapy on behavior profile and musical skills in young adults with severe autism. *J. Altern. Complement. Med.* 13, 709–712. doi: 10.1089/acm.2006.6334
- Carper, R., Solders, S., Treiber, J., Fishman, I., and Müller, R. A. (2015). Corticospinal tract anatomy and functional connectivity of primary motor cortex in autism. *J. Am. Acad. Child Adolesc. Psychiatry* 54, 859–867. doi: 10.1016/j.jaac.2015.07.007
- Catani, M., Jones, D., and ffytche, D. (2005). Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16. doi: 10.1002/ana.20319
- Chenausky, K., Kernbach, J., Norton, A., and Schlaug, G. (2017). White matter integrity and treatment-based change in speech performance in minimally verbal children with autism spectrum disorder. *Front. Hum. Neurosci.* 11:175. doi: 10.3389/fnhum.2017.00175
- Chenausky, K., Norton, A., Tager-Flusberg, H., and Schlaug, G. (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:e0164930. doi: 10.1371/journal.pone.0164930
- Christensen, D. L., Baio, J., Van Naarden Braun, K., Bilder, D., Charles, J., Constantino, J. N., et al. (2016). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2012. *MMWR Surveill. Summ.* 65, 1–23. doi: 10.15585/mmwr.ss6503a1
- Davis, B., and MacNeilage, P. (1995). The articulatory basis of babbling. *J. Speech Hear. Res.* 38, 1199–1211. doi: 10.1044/jshr.3806.1199
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., et al. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Friederici, A. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.0006.2011
- Gernsbacher, M., Sauer, E., Geye, H., Schweigert, E., and Hill Goldsmith, H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *J. Child Psychol. Psychiatry* 49, 43–50. doi: 10.1111/j.1469-7610.2007.01820.x
- Happé, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0
- Hardy, M., and LaGasse, A. (2013). Rhythm, movement, and autism: using rhythmic rehabilitation research as a model for autism. *Front. Integr. Neurosci.* 7:19. doi: 10.3389/fnint.2013.00019
- Iverson, J. M., and Fagan, M. K. (2004). Infant vocal-motor coordination: precursor to the gesture-speech system? *Child Dev.* 75, 1053–1066. doi: 10.1111/j.1467-8624.2004.00725.x
- Kasari, C., Brady, N., Lord, C., and Tager-Flusberg, H. (2013). Assessing the minimally verbal school-aged child with autism spectrum disorder. *Autism Res.* 6, 479–493. doi: 10.1002/aur.1334
- Kaufman, N. (1995). *The Kaufman Speech Praxis Test for Children*. Detroit, MI: Wayne State University Press.
- Kjelgaard, M., and Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Lang. Cogn. Process.* 16, 287–308. doi: 10.1080/01690960042000058
- Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314. doi: 10.1523/JNEUROSCI.4822-06.2007
- Lai, G., Pantazatos, S., Schneider, H., and Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain* 135, 961–975. doi: 10.1093/brain/awr335
- Leclercq, D., Duffau, H., Delamaire, C., Capelle, L., Gatignol, P., Ducros, M., et al. (2010). Comparison of diffusion tensor imaging tractography of language tracts and intraoperative subcortical stimulations. *J. Neurosurg.* 112, 5013–5511. doi: 10.3171/2009.8.JNS09558
- López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., and de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U S A* 110, 13168–13173. doi: 10.1073/pnas.1301696110
- Lord, C., Rutter, M., DiLavore, P., and Risi, S. (1999). *Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Services.
- Makris, N., and Pandya, D. (2009). The extreme capsule in humans and rethinking of the language circuitry. *Brain Struct. Funct.* 213, 343–358. doi: 10.1007/s00429-008-0199-8
- Meister, I., Boroojerdi, B., Foltys, H., Sparing, R., Huber, W., and Töpper, R. (2003). Motor cortex hand area and speech: implications for the development of language. *Neuropsychologia* 41, 401–406. doi: 10.1016/s0028-3932(02)00179-3
- Mullen, E. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service, Inc.
- Ozdemir, E., Norton, A., and Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *Neuroimage* 33, 628–635. doi: 10.1016/j.neuroimage.2006.07.013
- Plaisted, K. C. (2001). “Reduced generalization in autism: an alternative to weak central coherence,” in *The Development of Autism: Perspectives from Theory and Research*, eds J. A. Burack, T. Charman, N. Yirmiya and P. R. Zelazo (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 149–169.
- Rapin, I., Dunn, M., Allen, D., Stevens, M., and Fein, D. (2009). Subtypes of language disorders in school-age children with autism. *Dev. Neuropsychol.* 34, 66–84. doi: 10.1080/87565640802564648
- Rolheiser, T., Stamatakis, E., and Tyler, L. (2011). Dynamic processing in the human language system: synergy between the arcuate fascicle and the extreme capsule. *J. Neurosci.* 31, 16949–16957. doi: 10.1523/JNEUROSCI.2725-11.2011
- Rose, V., Trembath, D., Keen, D., and Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. *J. Intellect. Disabil. Res.* 60, 464–477. doi: 10.1111/jir.12284
- Saur, D., Kreher, B., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U S A* 105, 18035–18040. doi: 10.1073/pnas.0805234105
- Schlaug, G., Marchina, S., and Norton, A. (2009). Evidence for plasticity in white matter tracts of chronic aphasic patients undergoing intense intonation-based speech therapy. *Ann. N Y Acad. Sci.* 1169, 385–394. doi: 10.1111/j.1749-6632.2009.04587.x
- Schlaug, G., Norton, A., Marchina, S., Zipse, L., and Wan, C. (2010). From singing to speaking: facilitating recovery from nonfluent aphasia. *Future Neurol.* 5, 657–665. doi: 10.2217/fnl.10.44
- Tager-Flusberg, H., and Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Res.* 6, 468–478. doi: 10.1002/aur.1329
- Tager-Flusberg, H., Paul, R., and Lord, C. (2005). “Language and communication,” in *Handbook of Autism and Pervasive Developmental Disorder*, 3rd Edn, Vol. 1, eds F. Volkmar, R. Paul, A. Klin and D. Cohen (New York, NY: Wiley), 230–262.
- Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., et al. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum



- disorders. *J. Speech Lang. Hear. Res.* 52, 643–652. doi: 10.1044/1092-4388(2009/08-0136)
- Wan, C., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., 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:e25505. doi: 10.1371/journal.pone.0025505
- Wan, C., Zheng, X., Marchina, S., Norton, A., and Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain Lang.* 136, 1–7. doi: 10.1016/j.bandl.2014.03.011
- Zipse, L., Norton, A., Marchina, S., and Schlaug, G. (2012). When right is all that is left: plasticity of right-hemisphere tracts in a young aphasic patient. *Ann. N Y Acad. Sci.* 1252, 237–245. doi: 10.1111/j.1749-6632.2012.06454.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor currently co-hosts a Research Topic with one of the authors GS, and confirms the absence of any other collaboration. He states that the process met the standards of a fair and objective review.

*Copyright © 2017 Chenausky, Norton and Schlaug. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*