RESEARCH ARTICLE

The Effects of an Auditory Matching iPad App on Three Preschoolers' Echoic and Listener Responses

Lin Du¹ · Jeanne Speckman¹ · Megan Medina¹ · Michelle Cole-Hatchard¹

Published online: 23 February 2017

© Association for Behavior Analysis International 2017

Abstract We investigated the effects of an auditory match-to-sample protocol on three preschoolers' accurate echoics to 100 English words and advanced listener responses. The protocol was presented by using an iPad app Sounds the same: an app to target listening and speaking clearly. We used a combination of a multiple probe design (for echoic responses) and a delayed multiple probe design (for advanced listener literacy responses) with a time-lagged baseline across participants to test the effectiveness of the protocol. The three participants ranged from 4 to 5 years old and were all diagnosed as preschoolers with disabilities. They were taught to discriminate between positive and negative exemplars of progressively more difficult sounds, words, and phases by matching the sample stimulus to the matching exemplar. Our data show that the mastery of the intervention resulted in increases in the accuracy of the participants' articulation of their echoics, as well as their advanced listener repertoires as measured by the responses to spoken directions in the presence of visual distractors.

Keywords Auditory match-to-sample protocol · Echoics · Listener responses

Echoics occur early during children's development. Michael (1982) and Vargas (1986) referred to the echoic as a "duplic," a verbal response that has the formal similarity to its controlling verbal stimulus. Catania (1998) proposed that echoic behavior should not be defined by its acoustic correspondence,

but rather by correspondence of the phonetic units. Echoics are often confused with parroting, which does not require reinforcement by listeners of correspondence of phonetic units and cannot be counted as true verbal behavior (Skinner, 1957). Therefore, children who engage in non-functional repetitive parroting (echolalia) may not have a true echoic repertoire. Sundberg, Michael, Partington, and Sndberg (1996) argued that the duplication of speech sounds itself functions as automatic reinforcement for echoic behavior. Greer and Du (2015) and Greer and Ross (2008) further suggested social attention joining hear and say is the underlying conditioned reinforcement for echoic behavior.

Echoic behavior plays a critical role in child's early language acquisition. For a young child with a limited vocabulary, he acquires new words through echoics modeled by his teachers and parents. When his own production of speech resembles the model, the echoic is reinforced by this one-to-one correspondence. This correspondence then reinforces more future echoic responses, helps in building more advanced verbal responses such as tacts and mands, and eventually establishes his history of verbal behavior learning. Typically, developing children appear to acquire echoics effortlessly. One-year-old infants start to imitate sounds to get their parents' attention and at 2 years old, children are observed to echo words they hear from their parents independently (Learn the Signs. Act Early. <http://www.cdc.gov/ncbddd/actearly/index.html>). However, for individuals with language delays, echoic instruction has to be carefully engineered to provide them with ample opportunities to induce the unprompted vocalization and/or to improve the poor articulation. Lovaas (1977) stressed the importance of establishing children's "verbal imitations" that matched the adults' vocalization before teaching other verbal operants such as mands and tacts. Basic training that provides children with three to five echoic prompts at the onset of each training

✉ Lin Du
du@exchange.tc.columbia.edu

¹ Fred S. Keller School, Teachers College Columbia University, 1 Odell Plaza, Yonkers, NY 10701, USA

session is found to be helpful in increasing their articulation and even with independent speech (i.e., mands and tacts) (Kodak & Clements, 2009).

However, for children who have not yet emitted any true words or varying consonant-vowel combinations at all, we need to induce their first instance of vocal sounds. Tsiouri and Greer (2003), Tsiouri and Paul (2012), and Ross and Greer (2003) demonstrated fast and accurate responses from the rapid motor imitation task, which preceded a teacher's echoic prompt, helped to successfully induce echoic responses, and even led to independent mand and tact responses in their young participants who initially had no vocal communication skills. Another procedure, stimulus-stimulus pairing (Miguel, Carr, & Michael, 2002; Sundberg, Michael, Partington, & Sundberg, 1996; Smith, Michael, & Sundberg, 1996; Yoon & Bennett, 2000), has also been shown to be effective in promoting children's rate of initial vocal behavior, which serves as a building block for further shaping of vocalizations. Successive shaping (Skinner, 1953) gradually shapes the initial vocalizations to more accurate vocal responses that resemble more correspondence with their vocal models with differential reinforcement. This process also allows the contingencies to be shifted from automatic reinforcement to be under operant control and become functional (Bijou & Baer, 1965).

Yet, another important verbal behavior developmental cusp (Rosales-Ruiz & Baer, 1997), which enables one to learn at an accelerated rate or learn things they could not before, is the ability to respond to only the auditory properties of other's vocal directions or listener responses (Greer & Ross, 2008). Without listener responses, one cannot reliably discriminate between different vowel-consonant combinations and respond accordingly. Crystal (2006) pointed out the importance of phonemic awareness in children's language development in the induction of their joint attention or naming experience, a critical verbal behavior cusp that enables one to respond to a novel stimulus as a speaker with a tact and with a selection response as a listener upon hearing its name without direct instruction (i.e., Fiorile & Greer 2007; Greer, Stolfi, Chavez-Brown, & Rivera Valdez, 2005; Greer et al., 2007). Greer and Ross (2008) argued the importance of being a true listener and responding to the vowel-consonant properties of commands without any visual cues. Greer, Chavez-Brown, Nirgudkar, Stolfi, and Rivera-Valdes (2005) reported the identification and establishment of basic listener responses through intensive immersion in listener instruction with all eight participants, and the acquisition of the cusp resulted in their accelerated learning rates. Once children acquire basic listener responses, additional instruction introducing finer auditory discriminations may lead to more advanced listener responses, which allows them to follow a vocal direction with a simultaneously presented visual distractor.

Previous research shows that an effective intervention that targets both echoic and listener responses is the auditory match-to-sample (AM) protocol. This procedure has been found to be effective in the emergence of first instances of echoics, enhancement of pronunciation (Chavez-Brown, 2005), and induction of the listener component of naming (Speckman-Collins, Park, & Greer, 2007) with students who have language delays and autism spectrum disorders. In a recent paper, Choi, Greer, and Keohane (2015) reported that elementary students with autism spectrum disorders who mastered the AM protocol demonstrated the acquisition of conditioned reinforcement for correspondence between identical speech sounds, which resulted in improved phonemic awareness, advanced listener responses, articulation, and even full naming.

The materials used for the AM protocol have evolved given scientific and technological advances. Earlier studies (Chavez-Brown, 2005; Speckman-Collins, Park, & Greer, 2007) used BIGmack® buttons that allowed up to 20-s voice recording. However, these enabling devices were very bulky and pricey and the teacher needed to record a large number of them, with multiple negative exemplars as well, in order to complete one training session. Choi, Greer, and Keohane (2015) developed a newer version of the Adobe Flash® program, which could be played on a computer with softwares that supported Adobe Flash. The computer-based program greatly decreased the preparation time for educators as well as the instructional time for students. However, teaching very young students, especially those who have fine motor deficits, to use a computer mouse still presented some difficulty during the implementation of the program. In response to these technical challenges, we developed an iPad app *Sounds the same: an app to target listening and speaking clearly* (with the donated copyright from Dr. Choi). With this new app, even very young children and those with fine motor developmental challenges were able to respond to the instruction from the AM protocol by easily touching the screen with their fingers or with the use of a stylus. The purpose of the present study was to test the effects of this iPad app of AM protocol on clear echoics and advanced listener responses with young children with developmental disabilities.

Method

Participant

Three preschoolers (two males and one female) served as participants in this study. They attended a privately run, publicly funded preschool for children with and without developmental delays. The preschool was an accredited CABAS® school (www.cabasschools.org). Participant A was diagnosed with Rubinstein-Taybi syndrome. This is a genetic disorder

Table 1 Demographical information and standard test scores of three participants

	Participant A	Participant B	Participant C
Age	4 years	4 years 7 m	5 years
Gender	Male	Female	Male
Standard test scores	Preschool Language Scale –5 (March 18, 2014) Auditory comprehension SS 50, 1% Expressive Communication SS 63, 1% Total communication SS 53, 1%	Preschool Language Scale –5 (July 30, 2013) Auditory comprehension SS 50, 1% Expressive communication SS 50, 1% Total communication SS 50, 1% Developmental assessment of young children 2 (July 16, 2013) Receptive language SS < 50, 0.1%	Vineland-II Interview Ed (March 04, 2015) Communication SS 49
Related Verbal Behavior Cusps and Capabilities	Generalized matching Conditioned reinforcement of attending to adults' voices Generalized imitation	Generalized matching Conditioned reinforcement of attending to adults' voices Generalized imitation Echoic to mand Echoic to tact	Generalized matching Conditioned reinforcement of attending to adults' voices Generalized imitation Echoic to mand

The Preschool Language Scale-5 is a developmental language skill assessment (Zimmerman, Steiner, & Pond, 2011). *The Developmental Assessment of Young Children (DYAC-2)* provides an initial assessment for determination of development delays (Voress & Maddox, 2013). *The Vineland-II Adaptive Behavior Scale* is an adaptive skill measurement (Sparrow, Cicchetti, & Balla, 2005)

characterized by short stature, facial abnormalities, and moderate to severe intellectual disabilities. Participants B and C were identified as children with autism spectrum disorders by the psychologists from the referring school districts.

The participants' verbal repertoires were assessed through the use of *CABAS International Curriculum and Inventory of Repertoires for Children from Preschool through Kindergarten* (C-PIRK®) (Greer, 2014; Waddington & Reed, 2009) and *Verbal Behavior Development Assessment* (VBDA) (Greer & Ross, 2008) by professionals who were calibrated to a standard for doing the assessments. The C-PIRK served as both a criterion-referenced assessment and a curriculum for children from preschool to kindergarten, and it included the following four major domains: academic literacy, community of reinforcers, self-management, and physical development. The VBDA was an assessment of children's development of listening, speaking, reading, writing, and other social functions of language.

At the onset of the study, participant A had a beginner listener repertoire and could attend to stimuli presented during instruction and match objects and pictures upon request. However, he was unable to use vocal or non-vocal approach to communicate with others. Participants B and C both had beginning listener and beginning speaker repertoires and can match objects and pictures upon request as well as emitting echoic to mands and echoic to tacts using single words. See Table 1 for their chronological ages and related standardized testing scores.

These three participants were chosen for the study because they had acquired the prerequisite skills such as generalized

matching for pictures, sit still, and basic gross motor imitation but had little success with previous intervention efforts in improving their inexact vocal echoics and low levels of correct listener responses to vocal directions prior to the study.

Setting

The study took place inside the participants' classrooms. The rest of the class continued to receive individualized instruction with their teachers or attended related services. During the echoic response probes, the teacher sat face to face with the participant and presented 100 English words vocally without demonstrating any visual prompts. During the advanced listener response probes, the teacher and the participant also sat facing each other. The teacher presented the one-step vocal directions (i.e., touch ear, roll arms) simultaneously with the visual distractors (i.e., touch eyes, touch nose). During the intervention, the teacher presented the AM protocol by using an iPad, which was positioned on a slightly tilted stand on the table for the participant to see all three buttons on the screen. A portable speaker connected with the iPad helped to amplify the volume and made sure the participants could hear the auditory stimuli clearly.

Dependent Variables

We measured the number of correct echoics of 100 English words (see Table 2) and the number of correct responses in the advanced listener response assessment (see Table 3) (Choi, Greer, & Keohane, 2015) during the pre- and post-

Table 2 The 100 English words probe list (Choi, Greer, & Keohane, 2015) with phonetic transcription (Small, 2005)

Word	Phonetic transcription	No.	Word	Phonetic transcription	No.	Word	Phonetic transcription	No.
About	/ ba t/	2	Impossible	/ m pas b l/	4	Teacher	/ tit /	2
Again	/ gɛn/	2	Independent	/ nd pɛnd nt/	4	Terrible	/ tɛr b l/	3
Almost	/ lmo st/	2	Into	/ ntʊ/	2	That's	/ ðæts/	1
Also	/ lso /	2	Its	/ ts/	1	Their	/ ðe r/	1
Always	/ lwe z/	2	Journal	/ d rn l/	2	Then	/ ðɛn/	1
Another	/ n ð /	3	Laugh	/ læf/	1	Through	/ θru/	1
Anyone	/ ɛ ni w n/	3	Let's	/ lɛts/	1	Trouble	/ tr b l/	2
Are	/ ar/	1	Lovable	/ l v b l/	3	Unhappiness	/ n hæpi/	4
Ask	/ æsk/	1	Made	/ me d/	1	Until	/ n t l/	2
Beautiful	/ bjut f l/	3	Member	/ mɛmb /	2	Usually	/ ju w li/	3
Because	/ b k z/	2	Mine	/ ma n/	1	Vacation	/ ve ke n/	3
Before	/ b f r/	2	Myself	/ ma sɛlf/	2	Very	/ vɛri/	2
Buy	/ ba /	1	Neighbor	/ ne b /	2	Want	/ want/	1
Can't	/ kænt/	1	New	/ nu/	1	Was	/ w z/	1
City	/ s t i/	2	News	/ nuz/	1	Wave	/ we v/	1
Community	/ k mjun ti/	4	No	/ no /	1	We're	/ w r/	1
Confusion	/ k n fju n/	3	Off	/ f/	1	Wear	/ we r/	1
Could	/ k d/	1	One	/ w n/	1	Weather	/ wɛð /	2
Countries	/ k ntriz/	2	Our	/ a r/	1	Went	/ wɛnt/	1
Didn't	/ d dnt/	2	Owl	/ a l/	1	Were	/ w r/	1
Discover	/ d sk v /	3	People	/ pip l/	2	Who	/ hu/	1
Doesn't	/ d z nt/	2	Prettier	/ pr ti /	3	Whole	/ ho l/	1
Don't	/ do nt/	1	Prettiest	/ pr ti ɛst/	3	Winner	/ w n /	2
Enough	/ n f/	2	Pretty	/ pr ti/	2	With	/ w θ/	1
Especially	/ spɛ li/	4	Probably	/ prab bli/	3	Won	/ w n/	1
Everybody	/ ɛ vri b di/	4	Question	/ kwɛst n/	2	Won't	/ wo nt/	1
Everything	/ ɛ vri θ ɪ/	3	Really	/ ril li/	2	Wouldn't	/ w dnt/	2
Except	/ k sɛpt/	2	Recycle	/ ri sa k l/	3	Write	/ ra t/	1
Exciting	/ ɛk sa t ɪ/	3	Right	/ ra t/	1	Writing	/ ra t ɪ/	2
Favorite	/ fe vr t/	2	Said	/ sɛd/	1	Yell	/ jɛl/	1
Friendly	/ frɛndli/	2	School	/ skul/	1	Young	/ j ɪ/	1
General	/ dɛ n r l/	3	Sister	/ s st /	1	You're	/ j r/	1
Getting	/ gɛt ɪ/	2	Something	/ s m θ ɪ/	2			
Hopeless	/ ho pl s/	2	Sometimes	/ s m ta mz/	2			

No. indicates the number of syllables in each word

intervention probes. No consequences were provided for the participants' responses during the probe sessions (Table 4).

Past research on AM protocol typically measured full echoic, partial echoic, and incorrect echoic responses (Chavez-Brown, 2005; Choi, Greer, & Keohane, 2015; Speckman-Collins, Park, & Greer, 2007). A full echoic response was defined as the participant emitting the vocal response with a point-to-point correspondence with the vocal antecedent presented by the teacher within 3 s (i.e., "about" in response to "about"). A partial echoic response was defined as the participant emitting the vocal response with only a partial

correspondence with the vocal antecedent presented by the teacher within 3 s (i.e., "bou" in response to "about"). An incorrect response occurred when the participant failed to respond within 3 s or emitted a vocal response that did not bear the point-to-point correspondence with the teacher's vocal antecedent (i.e., "dada" in response to "about").

In the current study, we adopted the international phonemic alphabet (IPA) (Small, 2005) and converted each English word into its phonetic transcription with an exact number of syllables. There was a total of 186 syllables in the 100 English words presented during the pre- and post-intervention echoic

Table 3 Advanced listener literacy probe (vocal directions and visual distracters used during listener literacy probe sessions) (Choi, Greer, & Keohane, 2015)

	Vocal directions	Visual distracters
1	Touch nose	Touch mouth
2	Touch ear	Touch eyes
3	Clap hands	Stand up
4	Blow kiss	Roll arms
5	Roll arms	Touch nose
6	Touch eyes	Touch ear
7	Touch mouth	Clap hands
8	Stand up	Blow kiss
9	Stomp feet	Touch head
10	Tap lap	Touch belly
11	Touch arm	Touch feet
12	Touch table	Touch knees
13	Touch head	Touch table
14	Touch belly	Touch arm
15	Touch feet	Tap lap
16	Touch knee	Stomp feet
17	Wave hands	Touch elbow
18	Raise arms	Touch shoulder
19	Touch elbow	Raise arms
20	Touch shoulder	Wave hands

probes. The breakdown of an entire word into small units allowed for accurate evaluation of the participants' echoic behavior. For example, when the teacher presented the word "about / ba t/" and the participant responded with the echoic

"/ ba t/" within 3 s, the teacher then recorded a plus (+) for the correctly echoed first syllable "/" and a minus (−) for the second syllable "/ba t/" because of the lack of point-to-point correspondence. After presenting all 100 words, the teacher summarized the total number of syllables echoed correctly by the participant. For the purpose of comparing the two data collection systems, the teacher also totaled the number of full, partial, and incorrect echoes.

The second dependent variable was the number of correct listener responses emitted by the participants during pre- and post-intervention probes. During the probes, the teacher presented the participant with a one-step vocal direction together with a visual distractor. For example, the teacher asked the participant to "touch nose" while she touched her mouth. If the participant only followed the teacher's vocal direction and touched his nose, it was counted as a correct listener response and a plus (+) was recorded. If the participant responded to the visual cue from the teacher and touched his mouth, it was counted as an incorrect listener response and a minus (−) was recorded. If the participant responded to both the visual and auditory antecedents and touched his nose and mouth, it was still counted as an incorrect listener response and a minus (−) was recorded. In addition, if the participant failed to respond within 3 s after the teacher presented the antecedent, it was counted as incorrect and a minus (−) was recorded.

Material

The AM intervention was presented through an iPad app *Sounds the same: an app to target listening and speaking clearly*. This app was launched by the Fred S. Keller School

Table 4 Materials used in basic and advanced auditory matching phases (Choi, Greer, & Keohane, 2015)

Basic training phases		Advanced training phases	
1	Match sound using no-sound as non-exemplar	1	Match words with end rhyme
2	Match sound using white sound as non-exemplar	2	Match words with last syllable rhyme
3	Match sound using other sounds as non-exemplar	3	Match words with double rhyme
4	Match word using sound as non-exemplar	4	Match words with triple rhyme
5	Match word using another word as non-exemplar (set 1)	5	Match words with beginning rhyme
6	Match word using another word as non-exemplar (set 2)	6	Match words with first syllable rhyme
7	Match word using another word as non-exemplar (set 3)	7	Match phrases with one different word out of two words
8	Match word using word with rhyming end as non-exemplar	8	Match phrases with one different word out of three words
		9	Match phrases with one different word in a sentence
		10	4-syllable and uncommon words (set 1)
		11	4-syllable and uncommon words (set 2)
		12	4-syllable and uncommon words (set 3)
		13	4-syllable and uncommon words (set 4)

and CABAS® and was developed based on the original *Adobe Flash®* Auditory Matching program (Choi, Greer, & Keohane, 2015). The intervention consisted of 22 phases, including one pre-intervention training phase, eight basic training phases, and 13 advanced training phases. The pre-intervention training phase was created to teach students the point-to matching topography. The basic section taught students to match target sounds/words versus no sound, white noise, and other sounds/words, and the advanced section taught students to match relatively longer and more complex words, short phrases, and simple sentences. As the protocol advanced, the auditory discriminations between the positive and negative exemplars became finer. The AM app recorded data automatically during the session and provided built-in automatic reinforcement for correct responses as well as correction procedure for incorrect responses. The sequence of the presentation was programmed to randomize in order to minimize the possibility of possible side biases.

Design

We used a combination of a multiple probe design (for echoic responses) and a delayed multiple probe design (for advanced listener responses) with a time-lagged baseline across participants to test the effectiveness of AM protocol via the iPad app. The sequence of the study was as follows: (1) the first participant was probed on echoics for 100 English words and advanced listener responses; (2) the first participant received AM protocol intervention; (3) once the previous participant demonstrated criterion level responding in the intervention, and the same steps were repeated for the next participant (the second and third participant received one additional pre-

intervention probe on echoic responses immediately prior to their entry to the intervention); and (4) after mastery of every four training phases during intervention, the teacher conducted a probe for the echoic and advanced listener response in the exact same manner as in the pre-intervention probes.

Procedure

Pre- and Post-Intervention Probes The teacher assessed the participants on their (1) accurate echoics to 100 English words and (2) advanced listener responses. During the echoic probe, the teacher followed the alphabet sequence of probe list and presented the 100 English words to the participant. The participant was given up to 3 s to respond. No consequences were provided for the responses. That is, no reinforcement was provided for correct responses and no corrections for incorrect ones. However, the participant was reinforced for their attending and participation. At the end of the probe, the teacher totaled the number of all correctly echoed syllables and presented on the participant's graph. Additionally, the total number of full, partial, and incorrect echoics were recorded and displayed out of 100 opportunities.

During the advanced listener response probe, the teacher presented 20 one-step vocal directions together with the modeling of visual distractors. The participant was not provided with any consequences for the responses.

Intervention During the AM intervention, a pre-intervention training phase was introduced before the start of the basic and advanced training phases. In this phase, three buttons were presented on the iPad screen, one on top and two at the bottom. Each of the three buttons had a cartooned picture (i.e. pig,

Fig. 1 Screenshot of a pre-intervention training phase

cow, and plane) (see Fig. 1 for an example), and its correspondent featured sound. The teacher pressed the top button (i.e., dog) followed by the two buttons on the bottom (i.e., dog and thunder) and returned to the top button (i.e., dog), then asked the participant to “match.” There was no sequence requirement in pressing the two buttons at the bottom; either clockwise or counter-clockwise was acceptable. It was suggested that both directions be employed during presentations to prevent potential side biases in the student’s responses. The participant was required to press one of two buttons at the bottom that resembled the identical picture and sound as the top one (i.e., dog) within 3 s. The teacher or the participant then clicked the button “submit” on the top right-hand corner of the screen to check whether the response was correct or not.

If the participant selected the button with the matching sound, it was counted as correct and a firework animation popped up in the center of the screen together with applause cheering sounds. The teacher also provided the participant with vocal praise or any pre-determined reinforcers. If the participant (1) pressed the top button, (2) selected a button with a different auditory stimulus, or (3) did not respond within 3 s, it was counted as an incorrect response. After the participant submitted an incorrect response, the app immediately prompted the participant to “try again” (shown in text in the center of the screen). The app also provided a model demonstration of the correct response by highlighting the two buttons with the matching pictures with the sounds. The teacher then repeated the presentation of the antecedent again and gave the participant another independent opportunity to respond. These steps were repeated until the participant pressed the correct matching button. The teacher provided a partial physical prompt to the participant at the third attempt during correction

procedure. The procedure remained the same as the participant progressed from phase to phase.

Once the participant met criterion (90% accuracy for two consecutive sessions or 100% accuracy for one session) in the pre-intervention training phase, he moved on to the basic training phases followed by the advanced training phases. In the basic and advanced training phases, the protocol interface remained largely the same, except that all three buttons appeared identical (i.e., three radios, three smiley faces, or three solid color circles) (see Fig. 2 for an example). In another word, there were no more visual prompts provided for the matching responses and the participant was required to respond solely to the auditory stimuli. The participant was required to press the button to match the target auditory stimulus with the identical auditory stimulus. The presentation of the instruction was identical to the pre-intervention training phase. The built-in automatic reinforcement and correction prompt remained the same in the basic and advanced training phases too. Once criteria were achieved for four training phases, a post-intervention probe was conducted to assess the participant’s potential progress in his articulation and advanced listener responses.

Interobserver Agreement

Interobserver agreement (IOA) was conducted by an independent observer taking data simultaneously with the teacher during the probe and intervention sessions. *Teacher Performance Rate and Accuracy Scale* (TPRA) measures were conducted by teacher mentors and classroom supervisors to evaluate the accuracy of the teachers’ performances in delivering this protocol. Once 100% accuracy was achieved, the teacher was

Fig. 2 Screenshot of a basic/advanced training phase

allowed to run the protocol independently with the student. The teacher mentors or the supervisors continued to conduct TPRAs to ensure the consistency and reliability in the teachers' presentation of the protocol. During the pre- and post-intervention probes for echoics and advanced listener responses, the teacher mentors or classroom supervisors conducted the TPRAs independently and compared the recorded data with the teachers'. IOA was calculated by dividing the number of point-to-point agreements and disagreements by the total numbers of agreements plus disagreements and multiplying by 100%. For participant A, IOA was collected for 83% of the echoic probe sessions with a mean of 99% (range, 98 to 100%), and 100% IOA was collected for both of the advanced listener response probe sessions. For participant B, IOA was collected for 57% of the echoic probe sessions with a mean of 98% (range, 97 to 100%), and 100% IOA was collected for both of the advanced listener response probe sessions. For participant C, 100% IOA was collected for all of the echoic probes and both of the advanced listener response probes. During AM intervention sessions, the teachers compared the data recorded on their data collection sheets with the

ones automatically collected by the app. IOA was 100% across all sessions during intervention sessions for all the participants.

Results

Figure 3 shows the three participants' number of correct response to the advanced listener response probe sessions. Participant A emitted nine correct responses in the pre-intervention probe and increased to 15 during the post-intervention probe. Participant B emitted six correct responses in the pre-intervention probe and increased to 14 during the post-intervention probe. Participant C emitted five correct responses in the pre-intervention probe and increased to 17 during the post-intervention probe.

Figure 4 shows the number of correct echoed syllables in the 100 English words emitted by the three participants during pre- and post-intervention probes. Prior to the intervention, all three participants emitted a minimal number of correct echoics to the English syllables. However, after the mastery of the

Fig. 3 Number of correct responses emitted by the participants during advanced listener literacy probes

Fig. 4 Number of syllables echoed correctly to 100 English words before and after mastery of every four auditory matching intervention phases

intervention, all emitted significantly higher correct responses in their echoes during the post-intervention probes. Participant A emitted two correct echoes in the pre-intervention probe and increased to 17 during the post-intervention probe. Participant B emitted two and three correct echoes in the pre-intervention probes and increased to 103 during the post-intervention probe. Participant C emitted 19 and 18 correct correct echoes in pre-intervention probes and increased to 162 during the post-intervention probe.

Figure 5 shows the number of full, partial, and incorrect echoes to the 100 English words emitted by the three participants during pre- and post-intervention probes. Participant A's full echoes increased from 1 to 11 words while his partial echoes increased from 1 to 4 words in the post-intervention probe. Participant B's full echoes increased from 1 to 51 words while his partial echoes increased from 1 to 30 words in the post-intervention probe. Participant C's full echoes increased from 8 to 81 words while his partial echoes increased from 6 to 18 words in the post-intervention probe.

Figure 6 shows the data collected during the intervention sessions. Participant A required 68 sessions to complete the AM protocol. Participant B mastered the training within 59

sessions. Participant C required 134 sessions before his mastery of all sessions in the AM protocol.

Discussion

Our results show that the acquisition of generalized auditory matching of words resulted in all three participants' increased accuracy in their articulation and advanced listener responses. Thus, our findings were in line with previous research in that the auditory matching protocol was effective in improving children's pronunciation as well as their listener responses (Chavez-Brown, 2005; Choi, Greer, & Keohane, 2015; Speckman-Collins, Park, & Greer, 2007). These participants for whom vocalizations were limited and articulation was flawed improved performance in these areas.

Our results from two out of three participants show clear effects in their improvement of accuracy of echoic responses while participant A showed relatively modest gains to the other two participants. He increased the number of syllables echoed correctly from 2 to 17 out of 186 compared. However, the magnitude of this modest growth might be a function of

Fig. 5 Number of full and partial echoes to 100 English words before and after mastery of every four auditory matching intervention phases

his rare genetic disorder (*Rubinstein-Taybi syndrome*). Severe language delays are one of the characteristics associated with this syndrome.

The current study has several educational implications for practitioners working with students with similar challenges and verbal behavior levels as the participants from the current study. First, before the AM intervention, visual stimuli were more likely to evoke responses than auditory stimuli for all participants. Following the AM intervention, the opposite was true, and this phonemic stimulus control was only achieved following the mastery of the AM protocol. Advanced listener responses represent one's ability to accurately respond to the acoustic properties of speech alone (even in the presence of competing visual stimuli), so we can say that the participants' advanced listener responses improved.

Second, participants who needed the most number of intervention sessions to achieve mastery criterion would also acquire the longest conditioning history, and this might attribute to the gains in their listener and speaker repertoires. In the current study, we noticed that participant C required almost double the number of sessions (an average of 6) compared to

three sessions with the other participants. However, he showed the most substantial increases in both his articulation and advanced listener responses compared to the other two participants. This supported the conditioned reinforcement theory suggested by Choi, Greer & Keohane (2015) that the auditory matching protocol helped to establish the conditioned reinforcement for speech sound combinations for the participants.

Third, the data collection system used in the current study captured the participants' more nuanced improvements compared to the previous studies that measured the emission of full echoes, partial echoes, and incorrect responses (Chavez-Brown, 2005; Choi, Greer, & Keohane, 2015; Speckman-Collins, Park, & Greer, 2007). We used the same 100 English words in the pre- and post-intervention probes, but each word was broken down into small units with its phonemic transcriptions. This updated IPA measurement has been particularly helpful for participants who demonstrated moderate gains in their echoes of syllables but were not yet at the stage of mastery of the full echoes. For example, an echoic “/bju/” in response to the word “beautiful” would be coded as

Fig. 6 Correct responses emitted by the three participants during auditory matching intervention sessions

a “partial echoic,” and a largely improved echoic response /’bjut / would still be given the same coding. That is, this type of “within word” progress would hardly be detected under the original data collection system, implying no progress from the participant. However, with our new measurement, the first echoic response would receive a coding for one out of three syllables correct and the second one for two of three correct. The student’s progress was not only immediately apparent and also mathematically comparable to his previous responses.

Fourth, the current study focused on a matching task with acoustic stimuli and required that participants responded as listeners by physically selecting the correct matching sound. The effects of this intervention, however, were seen on both listener and speaker responses. The effects of the AM intervention on the participants’ advanced listener responses were likely due to the conditioning of auditory stimuli as reinforcers. This would mean the participants were now more likely to observe these stimuli (Choi, Greer, & Keohane, 2015; Speckman-Collins, Park, & Greer, 2007). Participants’ speaker responses were positively affected as well, likely due to the same reason. It is noteworthy that this intervention was

chosen for these participants because traditional echoic teaching was not functioning to increase echoic behavior at an efficient rate. Additionally, the benefits of the AM intervention were seen across two repertoires which support the efficiency of the intervention for clients/students and practitioners. A limitation to this study is that we do not have the original echoic teaching data representing the ineffectiveness of this procedure. Future research should investigate and report on these data.

Last but not least, the AM program has improved substantially from its original format to be more user friendly. Initially, the AM protocol was implemented by using BIGmack® buttons, which was an expensive augmentative communication device that could only record one message (up to 20 s) at a time. To complete one intervention phase, it would need at least eight such buttons: two sets of four targeted sounds/words. Once the student proceeded from one phase to another, the teacher was required to wipe out the previous recordings on the buttons and worked on new ones. This whole process could be very labor and time-consuming.

More importantly, the quality of the recordings could vary widely among users. The second generation of the AM protocol was developed with the Adobe Flash® program and could be played on a touchscreen computer or any other computer with a mouse. This increased the convenience of implementation of the protocol markedly and decreased the preparation time for teachers compared to the use of the BigMack® buttons. However, students who did not have a touchscreen computer in their classrooms had to learn this protocol through the utilization of a mouse. This seemingly simple step appeared to add to the complexity of the procedure as well as the delay in the student's response latency significantly. Fortunately, the abovementioned technical difficulties were all removed for the potential users with the current version of the AM protocol. It was developed as an iPad app. The app streamlined the protocol so that all students were responding to the same auditory stimuli. Students could simply touch the iPad screen to activate the recorded sounds/words. One's progress is not limited by the challenges of technology for young children. This synthesis of technology and education therefore "opened the door for potential far-reaching consequences" for these individuals (Twyman, 2011). It is not only accessible to more professionals but also to all parents whose children have poor articulation or delayed listener repertoires. Moreover, the app provides automatic data collection and summarizes the data in visual displays that help track progress.

One of the limitations in the study was the one pre-intervention probe done for advanced listener response repertoire for all the participants prior to the start of the study. It would be ideal to complete at least one more such probe sessions for each participant in order to better control the participants' instructional history and maturation. The effects of the study would be even more robust with multiple probes after the participants' mastery of every four intervention phases. In addition, future research should look into the transcription of spontaneous speech samples in addition to echoics of the contrived word list. It is not uncommon that many children's clarity of articulation of each word decreases as their speech gets progressively more spontaneous and complex. Therefore, future research should look into more flexible measures of functional intelligibility, especially when they are presented in context (i.e., Goldman & Fristoe, 2000).

Acknowledgements The authors would like to thank Leslie Morales for her assistance in translating the 100 English words into their phonetic transcriptions.

Thanks are given to Lin Du, Jeanne Speckman, Fred S. Keller School, Department of Health and Behavior Studies, Teachers College Columbia University; Megan Medina, Michelle Cole-Hatchard, Fred S. Keller School.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Funding This study did not receive funding from any grants.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- Bijou, S. W., & Baer, D. M. (1965). *Child development: vol. 2. Universal stage of infancy*. New York: Appleton-Century-Crofts.
- Catania, A. C. (1998). The taxonomy of verbal behavior. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 405–433). New York: Plenum.
- Chavez-Brown, M. (2005). The effects of the acquisition of a generalized word match-to-sample repertoire on the echoic repertoire under mand and tact conditions. Available from ProQuest Dissertations and Theses database (Doctoral Dissertation). Columbia University. (UMI No. 3159725).
- Choi, J., Greer, R. D., & Keohane, D. D. (2015). The effects of an auditory match-to-sample procedure on listener and echoic responses. *The Behavioral Developmental Bulletin*, 20, 186–206.
- Crystal, D. (2006). *How language works: how babies babble, words change meaning, and language lives or dies*. Woodstock, New York: Overlook Press.
- Fiorile, C. A., & Greer, R. D. (2007). The induction of naming in children with no echoic-to-tact responses as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, 23, 71–88.
- Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe test of articulation 2* (Second ed.). Circle Pines, MN: American Guidance Services.
- Greer, R. D. (2014). *International curriculum and inventory of repertoires for children from pre-school through kindergarten* (6th ed.). Yonkers, NY: Fred S. Keller School and CABAS®.
- Greer, R. D., Chavez-Brown, M., Nirgudkar, A. S., Stolfi, L., & Rivera-Valdes, C. (2005a). Acquisition of fluent listener responses and the educational advancement of young children with autism and severe language delays. *European Journal of Behavior Analysis*, 6(2), 88–126.
- Greer, R. D., & Du, L. (2015). Identification and establishment of reinforcers that make the development of complex social language possible. *International Journal of Behavior Analysis and Autism Disorder*, 1(1), 13–34.
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: inducing and expanding new verbal capabilities in children with language delays*. Boston: Person Education, Inc..
- Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdez, C. (2005b). The emergence of the listener to speaker component of naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, 21, 123–134.
- Greer, R. D., Stolfi, L., & Pistoljevic, N. (2007). Emergence of naming in preschoolers: a comparison of multiple and single exemplar instruction. *European Journal of Behavior Analysis*, 8, 119–131.

- Kodak, T., & Clements, A. (2009). Acquisition of mands and tacts with concurrent echoic training. *Journal of Applied Behavior Analysis*, 42(4), 839–843. doi:10.1901/jaba.2009.42-839.
- Learn the Signs. Act Early. (2016, June 5th). Developmental milestones [online materials]. Retrieved from <http://www.cdc.gov/ncbddd/actearly/milestones/index.html>
- Lovaas, O. I. (1977). *The autistic child: language development through behavior modification*. New York: Irvington Publishers, Inc..
- Michael, J. (1982). Skinner's elementary verbal relations: some new categories. *The Analysis of Verbal Behavior*, 1, 1–3.
- Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior*, 18, 3–13.
- Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: a developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, 30(3), 533–544.
- Ross, D. E., & Greer, R. D. (2003). Generalized imitation and the mand: inducing first instances of speech in young children with autism. *Research in Developmental Disabilities*, 24, 58–74.
- Skinner, B. F. (1953). Shaping and maintaining operant behavior. In B. F. Skinner (Ed.), *Science and human behavior* (pp. 91–106). New York: Free Press.
- Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group and the B. F. Skinner Foundation.
- Small, L. H. (2005). *Fundamentals of phonetics: a practical guide for students*. Boston: Pearson/Allyn and Bacon.
- Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *The Analysis of Verbal Behavior*, 13, 39–48.
- Sparrow, S., Cicchetti, D., & Balla, D. (2005). *Vineland adaptive behavior scales, second edition (Vineland-II)*. American Guidance Services Publishing.
- Speckman-Collins, J., Lee Park, H., & Greer, R. D. (2007). Generalized selection-based auditory matching and the emergence of the listener component of naming. *Journal of Early and Intensive Behavioral Intervention*, 4(2), 412–429.
- Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, 13, 21–37.
- Tsiouri, I., & Greer, R. D. (2003). Inducing vocal verbal behavior through rapid motor imitation training in young children with language delays. *Journal of Behavioral Education*, 12, 185–206.
- Tsiouri, I., & Paul, R. (2012). *Rapid motor imitation antecedent (RMIA) training manual, research edition: teaching preverbal children with ASD to talk* (1st ed.). Baltimore, MD: Brookes Publishing.
- Twyman, J. (2011). Emerging technologies and behavior cusps: a new era for behavior analysis? *European Journal of Behavior Analysis*, 12, 461–482.
- Vargas, E. A. (1986). Intraverbal behavior. In P. Chase & L. Parrott (Eds.), *Psychological aspects of language*. Springfield, Ill: Charles C. Thomas.
- Voress, J., & Maddox, T. (2013). *Developmental assessment of young children, 2nd edition (DYAC-2)*. Austin, Texas: PRO-ED.
- Waddington, E. M., & Reed, P. (2009). The impact of using the “preschool inventory of repertoires for kindergarten” (PIRK®) on school outcomes for children with autistic Spectrum disorders. *Research in Autism Spectrum Disorders*, 3, 809–827.
- Yoon, S. Y., & Bennett, G. (2000). Effects of a stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior*, 17, 75–88.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool language scale, fifth edition (PLS-5)*. San Antonio, TX: Pearson.