



Modeling Training of Child's Echoic Conversational Response for Students with Autism Spectrum Disorder: To Be a Good Listener

Natsumi Ishikawa^{1,2} · Mikimasa Omori^{1,3,4} · Jun-ichi Yamamoto⁵

Published online: 31 May 2018
© Association for Behavior Analysis International 2018

Abstract

Students with autism spectrum disorder (ASD) often have difficulties in responding to conversation with verbal language. These students often repeat what they hear, and their echoic behavior has a potentially communicative function. We define the echoic behavior when an individual repeats a peer's topic word with appropriate prosody within 3 s as the *child's echoic conversational response*. In this study, we examined the acquisition of the child's echoic conversational response skills and whether these skills could provide and generalize natural conversation for 4 students with ASD. During the training, students were instructed to imitate the topic word that the experimenter had used in the latest conversation. Students learned the child's echoic conversational response skills and improved their conversation skills. They even showed a slight generalization for nontraining materials through trainings and improvements in responding with new verbal responses. These findings suggested that expanding speakers' repertoires for students with ASD might facilitate improvement of natural conversation skills.

Keywords Students with autism spectrum disorder · Child's echoic conversational response · Conversation skills · Echoic · Intraverbal

Difficulties in Conversational Responding for Individuals with Autism Spectrum Disorder

Individuals with autism spectrum disorder (ASD) often exhibit delays in communication such as the inability to

sustain conversation and difficulties with social interaction (*Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition [DSM-V]; American Psychiatric Association [APA], 2010). Conversation skills of students with ASD are generally limited, and they often find it difficult to respond to conversational initiations (Jones & Schwartz, 2009; Nadig, Lee, Singh, Bosshart, & Ozonoff, 2010). Responding with conversational initiation includes verbal and nonverbal response produced by an individual in order to answer or clarify the initiation made by the speaker and maintain a joint activity with the speaker (Casenhiser, Binns, McGill, Morderer, & Shanker, 2015). Previous research has shown that students with ASD were not good at responding with expressive language (verbal responses) to conversation, whereas they were appropriately able to respond with nodding and shaking their heads (nonverbal responses) to conversation (Capps, Kehres, & Sigman, 1998). Namely, students with ASD often have difficulty responding to the conversation with an appropriate verbal response (Jones & Schwartz, 2009; Nadig et al., 2010). Therefore, teaching responding conversational skills may help these individuals to develop functional communication skills.

A portion of the data in this article was previously presented as part of the master's thesis of the first author, written in Japanese.

✉ Mikimasa Omori
mikkiomori@swu.ac.jp

- ¹ Department of Psychology, Graduate School of Human Relations, Keio University, Minato-ku, Tokyo, Japan
- ² Department of Child Psychiatry, The University of Tokyo Hospital, Bunkyo-ku, Tokyo, Japan
- ³ Department of Psychology, Faculty of Human and Social Sciences, Showa Women's University, Taishido, Setagaya-ku, Tokyo 154-8533, Japan
- ⁴ Research Fellow of the Japan Society for the Promotion of Science Chiyoda-ku, Tokyo, Japan
- ⁵ Department of Psychology, Faculty of Letters, Keio University, Minato-ku, Tokyo, Japan

Conversational Response with Modeling Training for Individuals with ASD

To teach conversational responding skills, both video and in vivo modeling trainings (Charlop-Christy, Le, & Freeman, 2000; Gena, Couloura, & Kymissis, 2005) and behavioral skills training (BST; Hood, Luczynski, & Mitteer, 2017; Peters & Thompson, 2015) are often used. In video modeling, participants usually observed videotaped conversation of two people, whereas participants observe a live model of conversation in in vivo modeling (Charlop-Christy et al., 2000). In typical BST, participants are instructed to model the response made by the trainer and rehearse the response by themselves. Although individuals with ASD learned the conversation skills of question asking (Charlop-Christy et al., 2000; Peters & Thompson, 2015), greeting (Hood et al., 2017), and affective behaviors of sympathy, appreciation, and disapproval (Gena et al., 2005), they rarely learn the varieties of verbal response during conversation (Charlop & Milstein, 1989).

Charlop and Milstein (1989) recorded the number of appropriate novel verbal responses after video-modeling training for three children with ASD. Their results showed that the occurrences of new verbal responses increased from 2.26% in baseline to 30.43% in posttraining. However, their results included the echolalic response in which the participants echoed the experimenter's vocalization. In order to expand verbal response varieties, we need to exclude the occurrences of echolalic response because it is thought *not* to facilitate language development for students with ASD (Sterponi & Shankey, 2014; Tager-Flusberg et al., 1990).

Responding with Echoic Behavior to Conversations and Training with Echoic Prompts

Whereas echolalia is a pathological repetition in which individuals repeat what they hear for one to many times (Skinner, 1957), echoic behavior shares point-to-point correspondence with the vocal-verbal behavior that evokes it (Skinner, 1957). Responding by echoic behavior permits speakers to react to it in other ways, especially in complicated conversation or directions to be followed. These responses are often emitted in asking for clarification or for expansion of what the speaker said, and acquiring echoic behavior may facilitate learning other verbal operants (Skinner, 1957). Namely, responding with echoic behavior should give the speaker another opportunity to initiate or react to the previous conversation. During conversation, one may also be required to respond with vocal responses related to, but not identical to, the verbal antecedents presented by another speaker (Greer & Ross, 2008). This verbal behavior is called intraverbal behavior, and it allows speakers to sustain the conversation and to respond with specific topics (Sundberg & Sundberg, 2011). This behavior is usually controlled by multiple complex verbal stimuli

(Eikeseth & Smith, 2013; Skinner, 1957; Sundberg & Sundberg, 2011).

Children with ASD acquire intraverbal responding to answer questions by using echoic prompts (Ingvarsson & Hollobaugh, 2010; Vedora, Meunier, & Mackay, 2009). In these studies, participants were taught to respond to the vocal and visual prompts when the experimenter asked participants questions verbally. The vocal prompts were then gradually faded, and the participants acquired the ability to respond to the questions with the appropriate verbal response as a result of the transfer of stimulus control. We therefore decided to present both echoic and visual prompts to teach the conversation skills through in vivo and video-modeling training. However, previous studies (Ingvarsson & Hollobaugh, 2010; Vedora et al., 2009) did not examine the effects of echoic prompts in responding conversation skills. We therefore examine the training effects of echoic prompts with video or in vivo modeling in teaching conversational responding to children with ASD.

Responding with a Part of Antecedences to Conversations as the Child's Echoic Conversational Response

With echoic prompts (Ingvarsson & Hollobaugh, 2010; Vedora et al., 2009) or modeling training (Charlop & Milstein, 1989), individuals with ASD can repeat what they hear and have the repertoire to listen to the conversation and to produce phrases. For example, if an individual said, "Very cold," with appropriate prosody, when he or she heard a peer say, "Today is really cold," echoic behavior can be considered as a communicative response. Thus, when an individual repeats, while responding, a peer's shortened phrase with appropriate prosody within 3 s, we define this echoic behavior as the *child's echoic conversational response*. The child's echoic conversational response is similar to echoic behavior in terms of having formal similarity between verbal stimuli and produced verbal responses (Skinner, 1957). However, the child's echoic conversational response does not necessarily have a point-to-point correspondence to the verbal stimuli, although responses with the child's echoic conversational response would show partial correspondence. Namely, the child's echoic conversational response would be a kind of intraverbal-echoic (Bondy, Tincani, & Frost, 2004) response in that the individual responded ("Very cold.") with a part of the verbal antecedents of another speaker ("Today is really cold."). It is necessary to respond with topic-related words presented by another speaker. In other words, the verbal response of the child's echoic conversational response is contingent on the verbal stimuli. We believe that by using this technique, individuals with ASD can expand their diversity of responses in natural conversation and exhibit generalized conversational skills.

Objectives

In this study, we examined two main questions regarding the child's echoic conversational response. First, we examined whether four students with ASD could acquire child's echoic conversational response skills and increase appropriate conversational responses through in vivo and video-modeling procedures. We prepared 20 conventional conversation scripts, of which half were used as training materials. With the other half, we evaluated whether our participants could generalize conversational skills to a new script and a new individual. Second, we asked other people to evaluate the social validity of the participants' conversational skills with 5-point Likert scales. If our participants improved their conversational skills by acquiring child's echoic conversational response skills, the scores from the Likert scales in the probe phase would be higher than that in the baseline phase.

Method

Participants

Four students diagnosed with ASD participated in the present study. Informed consent was obtained from both students and their parents. For students with ASD, their diagnosis was provided individually by a pediatric doctor using the DSM-IV-TR criteria (APA, 2000).

Table 1 displays participants' profiles with their names changed to protect their identities. At the time of the study, all students were enrolled in mainstream schools. Their mean age was 10 years 2 months ($SD \pm 1.84$). The participants' mean full-scale IQ (FSIQ) was assessed with Wechsler Intelligence Scales for Children, Third Edition (WISC-III; Wechsler, 1991/1998), or Wechsler Intelligence Scales for Children, Fourth Edition (WISC-IV; Wechsler, 2003/2010),

and was 91.00 ($SD \pm 14.26$). WISC-IV was used for only one student, Taku. The Japanese version of the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1986) was used to assess the students' autism severity, and the mean score of autism severity was 26.13 ($SD \pm 6.86$). For behavior screening, we implemented the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), and the mean total difficulties score was 18.00 ($SD \pm 2.94$). All participants were assessed using the Picture Vocabulary Test-Revised (PVT-R; Ueno, Nagoshi, & Konuki, 2008) to measure their verbal ages, and the mean verbal age was 8 years 7 months ($SD \pm 1.66$). The experimental procedures for participants Ryo and Yuta were conducted in a laboratory at Keio University, and they were trained with in vivo modeling. For participants Ken and Taku, we conducted procedures in rooms of their homes. We trained Ken and Taku with video modeling instead of using in vivo modeling because the assistant experimenter could not accompany the first author on a visit to their homes.

Apparatus and Stimuli

To record the participants' speech, all participants' conversational data were videotaped. A laptop computer (Apple MacBook Air 5.1) was used to present model videos. During the experiment, the first author began the conversation with a topic, and the participant was asked to respond and talk about the auditorily presented topic. Thus, we prepared 20 scripts, each of which contained two to four words including one topic word. We chose topics that the participants liked (e.g., trains, favorite meals, and cartoon characters).

Dependent Variables

Percentages of the child's Echoic Conversational Response Occurrence In this study, the child's echoic conversational

Table 1 Participant profile

Scale Participants	Chronological age (year; month)	WISC-III FSIQ	CARS Autism severity	SDQ Total difficulties scores	PVT-R Verbal age (year; month)
Ryo	8; 10	80	33	18	8; 10
Ken	10; 03	78	21	17	8; 09
Yuta	12; 09	99	31	15	9; 09
Taku	8; 10	107 ^a	19.5	22	5; 11
Mean	10; 02	91.00	26.13	18.00	8; 07
(SD)	(1.84)	(91.00)	(6.86)	(2.94)	(1.66)

Autism severity was measured using the Japanese version of the Childhood Autism Rating Scale (CARS; Schopler et al., 1986). The total difficulties score was measured using the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). Verbal age was measured using the Picture Vocabulary Test-Revised (PVT-R; Ueno et al., 2008)

^a Full-scale intelligence quotient (FSIQ) scores were measured by using WISC-III, but WISC-IV was used for Taku

response occurred if within 3 s the participant made a contingent vocal response on the topic word that the first author used to initiate the conversation. For example, when the first author initiated the conversation with “It is so cold today,” the participant was required to respond, “Very *cold*.” If the participant could respond with the topic word, we counted the response as an occurrence of the child’s echoic conversational response. If the participant responded without using the topic word, as in “Yes” or “I think so,” we did not count the response as an occurrence of the child’s echoic conversational response. We also counted as incorrect the participant’s response with the entire target sentence because we believe the response is not functional. After counting the occurrences, the percentages of the child’s echoic conversational response occurrences were calculated by dividing the number of occurrences by 10 (total trials in a session) and multiplying by 100.

Frequency of Appropriate Conversational Responses We also counted the mean frequency of appropriate conversational responses for baseline and probe sessions. In this study, a conversation always began with the experimenter’s initiation of the topic, and children were required to respond with vocal responses related to the topic. Therefore, we defined one conversational response as the child’s vocal response that contingently occurred within 3 s after the experimenter’s conversational initiation. For example, if the participant responded with the phrase “Very cold” in a trial within 3 s after the experimenter initiated with “It is cold today,” we counted an interaction as one appropriate conversational response. The experimenter continued to have conversations with the child by talking about related content on the topic when the child responded with the topic words. For example, the experimenter responded with related content on the topic, “I heard it is only 0°C today,” if the participant responded with “Very cold” in the latest conversation. The conversation finished when the child responded with nonverbal responses, such as nodding and conventional responses, or did not respond to the conversation for 10 s in a conversational turn. The conversation also finished if the child responded with unrelated topic words for three consecutive turns. Although responses with these interactions are not appropriate, we counted the child’s nonverbal responses after the experimenter’s initiation as one appropriate conversational response and also counted the child’s consecutive unrelated topic-words response for three turns as three conversational responses.

Frequency of Novel Words For the first baseline session, we counted the number of correctly vocalized words. We did not count the words that the participant had already spoken or repetitive and similar vocal responses within the sessions and trials. In the analysis, only words occurring for the first time in the session and the trials were included. In other words, we only counted a word as novel if the participant

had not spoken or modeled the word in all previous trials, including baseline, trainings, posttest, probes, and generalization phase. We did not count a word as novel if the word had already appeared across each session and trial. To evaluate vocal response diversities for pretraining word frequency, we subtracted the number of vocalized words in the last baseline session from the number in the first baseline session. For example, if the participant responded with the word “OK” in a trial of the first baseline session and “OK, thanks” in a trial of the last baseline session, and the participant had not spoken the word “thanks” in all trials of the previous baseline sessions, we only counted the word “thanks” as a novel word for pretraining. For posttraining word frequency, the number of vocalized words in the last probe session was also subtracted from the number in the last baseline session. For example, if the number of vocalized words were 10 in the first baseline session, 20 in the last baseline session, and 50 in the last probe session, the word frequency in pretraining was 10 and in posttraining was 30.

Proportion of Simple Echoing in Sessions By only responding with the topic word, the participants would learn nonfunctional speech through the child’s echoic conversational response training. We therefore calculated the proportion of simple echoing in the percentages of correct responses. We first added the number of correct responses (simple echoing) and novel word frequency in a session. For example, if the number of correct responses is 2 and the novel word frequency is 8, the total number of vocalizations becomes 10 in a session. We then divided 2, the number of simple echoing responses, by 10, the total number of vocalizations, to get 20% for the proportion of simple echoing in the percentages of correct responses.

Scores of Social Validity Our questionnaire contained eight items about conversation. On the basis of each rater, mean scores of each question for each participant were calculated (see the section on social validity for details).

Experimental Design

A multiple-baseline across-participants design (Kazdin, 2011) was used.

Procedure

We conducted preassessment, topic selection, baseline, modeling, and role-play training for the child’s echoic conversational response, probe, and generalization.

Preassessment We provided 10 phrases for the imitation task. By presenting five words and five short phrases, we could evaluate students’ abilities in imitation and listening skills.

We instructed students to respond to what they heard to assess the rule governing behavior skills. For example, if we say, “Say, ‘apple,’” or “Say, ‘Today is cold,’” students were required to respond by saying, “Apple,” or “Today is cold.” Each word or phrase was presented one by one, and all students correctly imitated five words and five phrases. After the imitation task, we had 3 min free conversation to evaluate their conversation skills. During conversation, however, all participants usually responded with one to three words (e.g., “No,” “I liked it.”). Sometimes they did not respond when the speaker asked them questions. In addition, even though they might respond to the speakers’ questions, participants usually gave stereotypic responses, such as “Yes,” “OK,” “Ah,” and so on.

Topic Selection After the preassessment, we asked parents to list up to 30 of their children’s favorite things (i.e., foods, cartoons, sports, video games, etc.), funny and warmhearted episodes, episodes of failure, and what their children wanted to be. For example, the experimenter initiated the conversation with “Hey, I want to eat steak for my dinner tonight [as a topic on food],” or “Yesterday, I watched a movie of [the name of a cartoon, as a topic on a cartoon].” Then, participants were required to respond to the conversation, such as with “OK,” as stereotypic responses and/or “Steak! Sounds good!” as the child’s echoic conversational response. For the baseline conversation, the first author then chose 20 topic words from the parents’ reports, including favorite sports, video games, cartoons, and a failure story.

Baseline Before starting the baseline conversation, the first author instructed the participant to talk about each topic word. After the participant agreed, the author and participant began to talk about the first topic. During the baseline phase, the experimenter always began to talk about the topic, and students were instructed to converse or respond during conversation. If participants responded with topic words or related words, we kept conversing on the topic. For example, when we talked about favorite sports, and the first author said, “Speaking of favorite sports, I like basketball,” the participant was required to respond with the topic word, for example, “Basketball? Good!” or related words, for example, “I like football better.” When the participant responded with topic or related words, the first author then sustained the conversation, such as by asking, “Do you like basketball?”

In case of comments, the first author maintained conversations until the child failed to emit any verbal response within 3 s, when participants indicated end of speech, and when participants made a comment unrelated to the conversation. For example, the experimenter and participant finished a conversation if the participant said, “OK,” accompanied by a high five. If the child provided some responses, the first author said, “Thank you for talking.” Then, the first author changed the topic of conversation. After finishing three or four

conversational sessions, until the occurrence of the child’s echoic conversational response stabilized, students were instructed to start modeling training for the child’s echoic conversational response.

Modeling and Role-Playing Training for the child’s Echoic Conversational Response Of 20 topic words, 10 were chosen for modeling training for the child’s echoic conversational response. The other 10 topic words were used for generalization tests. One session consisted of 10 trials with each topic word presented once.

First, experimenters instructed participants in the following manner: “In order to be a good listener, you need to listen carefully to what the speaker said, and you need to reply or imitate with the topic words in the conversation.” After this instruction, the experimenters began modeling training for the child’s echoic conversational response. In modeling training, Ryo and Yuta were trained with in vivo modeling, and Ken and Taku were trained with video modeling. In in vivo modeling, participants were instructed to watch the live-model setting carefully and the first author and the assistant experimenter, who was familiar with all participants, sat facing each other. The first author always initiated the conversation, such as by saying, “Speaking of dinner tonight, I want to eat steak,” and the assistant experimenter modeled as the participant was required to respond, that is, with the topic word from what she heard in the latest conversation, such as by saying, “Oh, steak?” Participants were instructed to observe how the assistant experimenter responded to the conversation. The video-modeling setting was identical to the setting of in vivo modeling except that participants observed the videotaped model presented on the computer. Immediately after participants observed one modeling situation for either setting, role-playing training was implemented.

In role-playing training for the child’s echoic conversational response, the first author initiated conversation on the same topic that the participant had just observed. The participant was required to respond with the topic words, just as the assistant experimenter did in the modeling training. If the participant could respond with the topic word, the first author gave him or her verbal praise, for instance, “What a great listener!” If the participant could not respond with the topic word, we gave him or her a verbal prompt, for instance, “Thank you for telling me, but listen more carefully.” We did not present the trial-specific topic word as a verbal prompt, such as “Say, ‘Soccer!’” After the prompt, the modeling scene was presented again, and the participant was instructed to respond to what the assistant experimenter said after listening to the initiated conversation. The modeling scene was presented until participants correctly identified what the assistant experimenter said. After the correct response, the participant returned to role-playing. During modeling and role-playing training for the child’s echoic conversational response phase, participants were trained until they

could respond with the topic word from the initiated conversation with more than 70% correct for three consecutive sessions. We only counted responses as correct if participants could respond with the topic words following the first presentation of the target topic. If they could not after the first presentation, we counted the trial as incorrect.

Probe In the probe phase, a conversation identical to that of the baseline phase was conducted; however, the first author talked about only the 10 training topics. The probe phase was conducted for two sessions.

Generalization In the generalization phase, the first author talked about 10 nontraining topics. The conversational setting was identical to that of the probe phase.

Social Validity

To determine the intervention's social validity (SV), 10 college students, unfamiliar with the study's purposes, from the Department of Psychology at Keio University, were shown a 3-min videotaped sample for each participant. The raters watched one video from the first baseline session and one video from the last probe session of the four participants. The raters were blind to the first or last phase of the video he or she observed. The experimenters displayed the videos randomly, and after viewing each segment, the rater completed an eight-item questionnaire for each segment in which they rated the child's conversational skills, such as motivation, turn taking, and prosody of conversation, on a 5-point Likert-type scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *neither agree nor disagree*, 4 = *agree*, 5 = *strongly agree*). Prior to rating the SV scores, we only instructed raters to score as if they responded to the conversation like the child in the video clip. The college students completed the evaluation in 40 min.

Reliability

Two independent observers, including the first author, evaluated whether the child's echoic conversational response occurred on 12 out of 44 total sessions (27%). Both observers watched as the participant responded during the conversation; the observers then independently evaluated whether the response included the topic word that the first author said in the latest conversation. The observers evaluated all conversations for each participant. Session-by-session interobserver agreement (IOA), calculated by dividing the number of agreements by the total number of agreements plus disagreements, and multiplying by 100, was used to determine interrater reliability. For all responses, IOAs were 90% for Ryo (range 80%–100%), 100% for Ken and Yuta (range 100%–100%, for both), and 95% for Taku (range 90%–100%), indicating a satisfactory agreement. Kappa (Cohen, 1968) was calculated

to measure interrater reliability for the child's echoic conversational response responses and was found to be .91 for all responses from all students.

Data Analysis

For the statistical analysis, we conducted paired *t*-tests to evaluate the conversational response skills (frequency of appropriate conversational responses: baseline average vs. probe average) and to evaluate the vocal response diversities during conversation (frequency of novel words: pretraining word frequency vs. posttraining word frequency). We also conducted another paired *t*-test (SV scores: baseline phase vs. probe phase) to evaluate conversational skills for each participant.

Results

Occurrence of the Child's Echoic Conversational Response

Figure 1 shows the percentage of the child's echoic conversational response occurrence for all participants. During the baseline phase, all participants scored around 20% for the training topics and 15% for the nontraining topics. However, after the modeling and role-playing training, all participants except Taku completed the training in three sessions; Taku needed four sessions. In the probe phase, the mean percentages of the child's echoic conversational response occurrences for training topics increased to 81.25% ($SE \pm 4.79$) and for nontraining topics to 60% ($SE \pm 0.00$) as a generalization test.

Frequency of Appropriate Conversational Responses and Frequency of Novel Words

Figure 2 shows the mean frequency of appropriate conversational responses in a trial (left panel) and the mean number of vocalized novel word frequency (right panel). Open shapes indicate the individual data and closed shapes indicate the mean data. On the left panel of Fig. 2, participants talked about related topics for 2.97 ($SE \pm 0.55$) conversational turns during baseline, whereas they continued to talk for 5.42 ($SE \pm 0.92$) conversational turns after training. A paired *t*-test revealed a significant increase in the appropriate conversational responses after the modeling training of the child's echoic conversational response, $t(3) = 3.40$, $p < .05$, $r = .89$.

On the right panel of Fig. 2, participants vocalized 55.75 ($SE \pm 8.36$) novel words in pretraining, whereas they spoke 135.00 ($SE \pm 17.54$) novel words in posttraining. For the statistical analysis, a paired *t*-test revealed a significant increase in the frequency of novel words after the modeling training of the child's echoic conversational response, $t(3) = 4.46$, $p < .05$, $r = .93$.

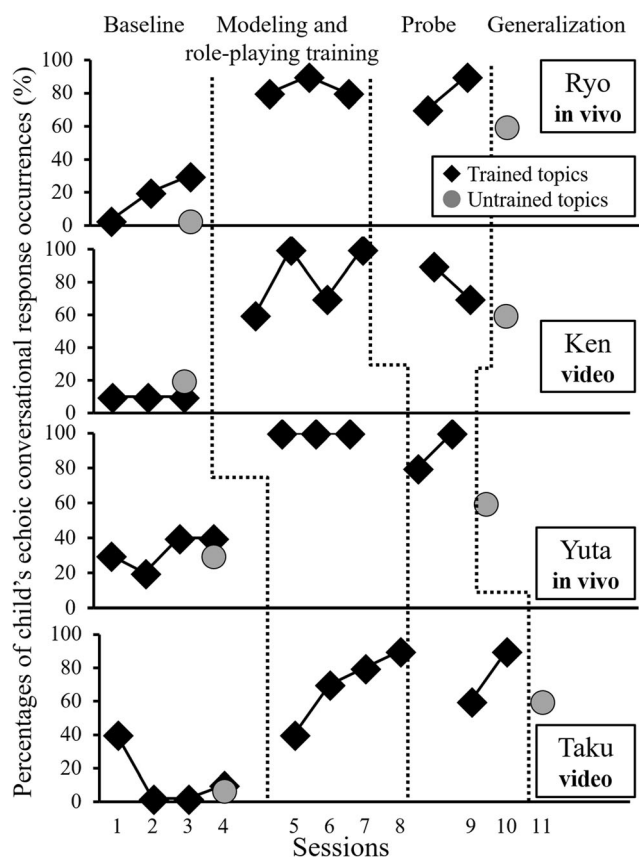
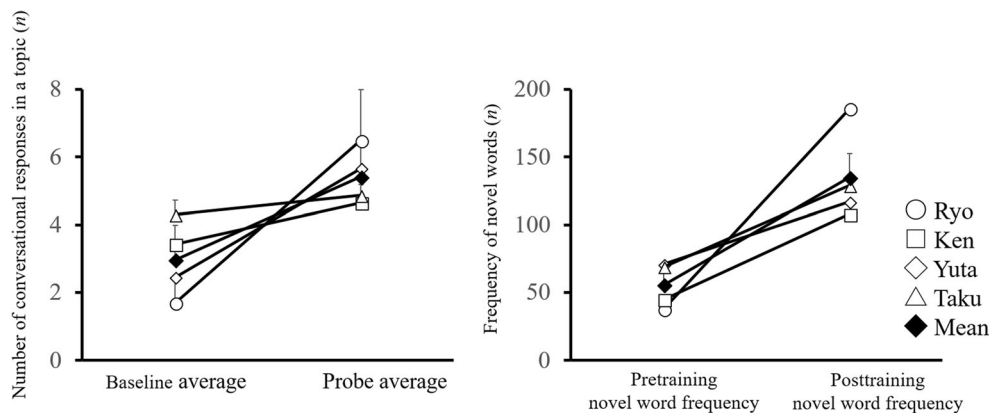


Fig. 1 Percentages of child's echoic conversational response occurrences

Proportion of Simple Echoing in Sessions

The mean proportion of simple echoing across all sessions was 13% ($SE \pm 0.02$) for Ryo, 21% ($SE \pm 0.04$) for Ken, 28% ($SE \pm 0.06$) for Yuta, and 14% ($SE \pm 0.04$) for Taku. The mean proportion of simple echoing for all participants was 19% ($SE \pm 0.03$, range 13%–28%), and we believe the score was low enough to report that the child's echoic conversational response training facilitated functional speech for children with ASD.

Fig. 2 Mean number of appropriate conversational responses (left panel) and mean number of the frequency of novel vocalized words (right panel). Open shapes show the individual data, and closed shapes indicate the mean data for both left and right panels



Scores of Social Validities (SV)

Table 2 shows scores of SVs for eight questions for each participant. During the baseline phase, their mean SV scores were 3.00 ($SE \pm 0.19$). After the training, all students improved in natural conversation skills, $t(3) = 3.52$, $p < .05$, $r = .90$, listening to conversation skills, $t(3) = 3.76$, $p < .05$, $r = .91$, fluent turn takings, $t(3) = 3.97$, $p < .05$, $r = .92$, engaging in the conversation, $t(3) = 4.44$, $p < .05$, $r = .93$, and using appropriate prosody, $t(3) = 3.98$, $p < .05$, $r = .92$. In addition, Ryo and Yuta improved in enjoying the conversation, Ryo: $t(9) = 15.41$, $p < .001$, $r = .99$; Yuta: $t(9) = 2.50$, $p < .05$, $r = .82$, being interested in the conversation, Ryo: $t(9) = 12.18$, $p < .001$, $r = .99$; Yuta: $t(9) = 3.01$, $p < .05$, $r = .87$, and being excited about the conversation, Ryo: $t(9) = 10.60$, $p < .001$, $r = .99$; Yuta: $t(9) = 2.87$, $p < .05$, $r = .86$.

Discussion

Learning the Child's Echoic Conversational Response by Modeling

We first examined whether students with ASD could acquire child's echoic conversational response skills through training with in vivo and video-modeling procedures. After the training, all students improved their percentages of the child's echoic conversational response occurrences for the 10 training topics. In addition, compared with the baseline phase (15%; $SE \pm 0.13$), they showed higher percentages of the child's echoic conversational response occurrences for 10 nontraining topics in the probe phase (60%; $SE \pm 0.00$). These findings replicate the results of previous studies (Charlop & Milstein, 1989; Charlop-Christy et al., 2000; Gena et al., 2005) that showed modeling training can improve conversational skills. In this study, Ken and Taku were trained with video modeling because they could not come to the lab room at Keio University and an assistant experimenter could not go to their

Table 2 Mean scores of social validities for conversational skills

Questionnaire item	Ryo		Ken		Yuta		Taku		Main (item)	
	Baseline	Probe	Baseline	Probe	Baseline	Probe	Baseline	Probe	Baseline	Probe
1. The participant is enjoying the conversation with the adult.	1.90 (0.18)	4.90**** (0.10)	4.40 (0.22)	4.60 (0.22)	2.20 (0.26)	3.10* (0.25)	4.10 (0.21)	4.00 (0.28)	3.15 (0.21)	4.15 (0.15)
2. The participant is talking naturally with the adult.	1.50 (0.22)	4.10**** (0.18)	3.70 (0.26)	4.10 (0.23)	2.50 (0.31)	3.60* (0.37)	2.90 (0.41)	3.80** (0.29)	2.65 (0.19)	3.90* (0.14)
3. The participant is good at listening to the adult's comments.	1.80 (0.20)	4.30**** (0.15)	3.80 (0.29)	4.30* (0.15)	3.40 (0.27)	4.20* (0.20)	2.70 (0.26)	3.90*** (0.28)	2.93 (0.17)	4.18* (0.10)
4. The participant is disinterested in the conversation with the adult (reversed item)	2.00 (0.21)	4.80**** (0.13)	4.40 (0.22)	4.50 (0.31)	2.80 (0.33)	3.70* (0.30)	4.30 (0.21)	4.00 (0.33)	3.38 (0.20)	4.25 (0.15)
5. The turn taking of the conversation is fluent.	2.00 (0.37)	4.30**** (0.21)	3.40 (0.27)	4.10* (0.28)	3.10 (0.31)	4.20* (0.20)	3.20 (0.25)	3.80* (0.25)	2.93 (0.17)	4.10* (0.12)
6. The participant is bored with the conversation (reversed item).	2.30 (0.26)	4.90**** (0.10)	4.20 (0.39)	4.60 (0.22)	3.10 (0.43)	3.90* (0.35)	4.50 (0.22)	4.10 (0.31)	3.53 (0.21)	4.38 (0.14)
7. The participant wants nothing to do with the conversation (reversed item).	2.20 (0.33)	3.80**** (0.29)	2.50 (0.31)	3.50** (0.27)	4.00 (0.30)	4.40 (0.34)	3.40 (0.31)	4.00* (0.15)	3.03 (0.19)	3.93* (0.14)
8. The participant is using the appropriate prosody in response to the adult's comments.	1.50 (0.22)	3.50**** (0.34)	3.70 (0.21)	3.90 (0.41)	1.70 (0.21)	3.00* (0.39)	2.90 (0.31)	3.90** (0.35)	2.45 (0.18)	3.58* (0.19)
Mean (individual)	1.90 (0.25)	4.33 (0.19)	3.76 (0.27)	4.20 (0.26)	2.85 (0.30)	3.76 (0.30)	3.50 (0.27)	3.94 (0.28)	3.00 (0.19)	4.06 (0.14)

Ten raters independently evaluated the conversation skills for each participant in random order. Standard error scores are indicated below each mean score

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$

homes along with first author. However, the percentages of the child's echoic conversational response occurrences were equally improved after in vivo modeling (from 25.71% to 85.00%) and video modeling (from 11.43% to 77.50%). Therefore, our results suggested that in vivo modeling was as effective for students with ASD to acquire conversation skills as video modeling.

Responding with the Child's Echoic Conversational Response to Sustain Natural Conversation with Others

To have natural conversation, an individual initiates a topic, responds to what others say, and by taking turns, expresses an adequate or appropriate vocal response. However, students with ASD are known to have difficulty responding with expressive language (Capps et al., 1998; Nadig et al., 2010) during an interactive conversation. Possibly, students with ASD are not good at choosing appropriate responses during conversation because of the many choices—other than a correct one—in natural conversation. Prior to the intervention, our participants were only able to sustain conversation for about three turns in a trial. During baseline sessions, Ryo usually responded with nodding and stereotypic vocal responses such as “Yes,” “No,” or “OK,” and therefore he could not initiate and sustain the conversation. Ken, Yuta, and Taku often responded with unrelated topics. During baseline conversation, Ken and Taku interrupted the other speaker's conversation, whereas Yuta did not continue to talk when he finished responding to the first conversation.

Based on the left panel of Fig. 2, however, all participants could improve their conversational responding skills and conversational reciprocity after acquiring child's echoic conversational response skills. All participants also improved their SV scores on turn taking in Table 2. Ishizuka and Yamamoto (2016) recently reported that adults' contingent vocal imitation increased the verbal interactions for children with ASD. They suggested that the child and the adult established a reciprocal verbal interaction and exchanged the roles of imitating and being imitated after the adult's contingent verbal imitation. In our study, the experimenter and the participant might exchange roles of initiating and being initiated in the conversation after the participant responded with the child's echoic conversational response, and this would lead to an increase in the appropriate conversational responses on related topics. These results suggest that responding with the child's echoic conversational response increases opportunities for students with ASD to sustain natural conversation with others. We believe that improvements in listening might facilitate their speaking because these skills are closely linked in conversation.

Learning Intraverbal-Echoic Behavior by Increasing Novel Word Frequency

Although our participants acquired child's echoic conversational response skills, applying intraverbal-echoic behavior as training might lead them to use stereotypic vocal responses during conversation. Charlop and Milstein (1989) reported that children with ASD increased new verbal responses from 2.26% in baseline to 30.43% in posttest after video-modeling training, but echolalic responses were included. However, the right panel of Fig. 2 shows that our participants spoke greater numbers of new vocabulary words after the training. The proportion of simple echoing was around 20% for all participants. These results suggest that our participants did not learn stereotypic behavior, but diverse responses occurred during conversation in two ways.

First, based on the left panel of Fig. 2, participants learned to respond to the conversation with related topic words by learning the child's echoic conversational response. If participants failed to respond with related topics, the experimenter terminated the conversation within three conversational turns. However, the number of appropriate conversational responses increased from 2.97 ($SE \pm 0.55$) during baseline to 5.42 ($SE \pm 0.92$) after training. Second, based on the right panel of Fig. 2, participants might learn to respond to the conversation with intraverbal behavior by increasing novel word frequency. Intraverbal behavior usually does not have point-to-point correspondence with the verbal stimuli that evokes it (Eikeseth & Smith, 2013; Skinner, 1957). If participants only learned echoic and/or intraverbal-echoic skills via modeling training, their novel word frequency would not increase to 135 words because echoic and intraverbal-echoic behavior share the whole or the part of vocal antecedences (Bondy et al., 2004). Therefore, our results expanded the previous findings that echoic behavior can be used as intraverbal responses and responses in natural conversation (Vedora et al., 2009), and modeling training for the child's echoic conversational response can increase opportunities of responding appropriately (Ingvarsson & Hollobaugh, 2010).

The Effectiveness of the Child's Echoic Conversational Responses for Individuals with ASD and Limited Expressive Vocabulary

On the basis of our results, Ryo and Yuta seemed to learn child's echoic conversational response skills more effectively than Ken and Taku. Although their FSIQ or total difficulties scores on SDQ did not differ, Ryo and Yuta had lower autism severity scores on CARS (mean 32.00) than Ken and Taku (mean 20.25). Table 2 shows that raters scored Ken and Taku as enjoying conversation more, as getting more excited, and as

being more interested during conversation with others in the baseline phase than Ryo and Yuta. Although Taku's verbal age was lowest, he spoke 88 words, and Ken spoke 43 words in the first baseline phase, whereas Ryo and Yuta spoke only 25 words. Both Ken and Taku had a greater expressive vocabulary and showed greater responsiveness to conversation. In other words, they talked freely rather than imitating topic words. Although further investigation is necessary, modeling training for child's echoic conversational response skills seems more effective for individuals with high autism severity scores, who show little interest in conversation with others, and who produce less expressive vocabulary.

Future Research for Triadic Conversational Interactions with the Child's Echoic Conversational Response

Based on the left panel of Fig. 2 and question 5 in Table 2, learning child's echoic conversational response skills can reduce problems of dyadic turn taking during conversation. Children with ASD first learned dyadic interactions between themselves and another individual; thereafter, they learned triadic interactions among themselves and two other individuals (Tomasello, 1995). However, learning triadic interactions is one of the most difficult skills for children or students with ASD (e.g., Mundy, Sigman, & Kasari, 1994). For example, children with ASD have difficulty acquiring the joint attention needed in a triadic interaction (Mundy & Stella, 2000). Conversation among three people requires more turn-taking skills than conversation between two people. Avenues for future research therefore include examining whether students with ASD can use child's echoic conversational response skills in triadic conversation or whether additional training to have natural conversation among three people needs to be developed.

Limitations

With such a small sample, we should have collected more data, especially on the baseline phase for both trained and untrained topics. Although participants learned the child's echoic conversational response, the training phases were not identical for all participants. Namely, some topics would have been stronger for some participants to respond to. Therefore, in further investigations, the same 10 topics should be selected for all participants in all conditions. In addition, to sustain natural conversation, individuals should learn how to initiate the conversation, whereas participants in this study only learned to respond to the conversation. Future research should also include modeling training for initiating the conversation.

Implications for Practice

- Child's echoic conversational response was effective in improving and generalizing conversational skills for children with autism spectrum disorder (ASD).
- Both in vivo and video-modeling training were effective to teach the child's echoic conversational response.
- By responding with the child's echoic conversational response, the frequency of conversational responses increased.
- By sustaining the natural conversation, participants produced novel words frequently.
- The child's echoic conversational response would be more effective for individuals with ASD and limited expressive vocabulary.

Funding This research was supported in part by a grant-in-aid from the Japan Society for the Promotion of Science (JSPS) Research Fellowship (Grant No. 26-3116) and the JSPS KAKENHI (Grant No. 26285213).

Compliance with Ethical Standards

Conflict of Interest All three authors report no conflicts of interest.

Human Subjects All procedures performed in studies involving human participants were in accordance with the ethical standards of the Keio University Institutional Review Board of the Faculty of Letters 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 and their parents.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)*. Washington, DC: Author.
- American Psychiatric Association. (2010). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.
- Bondy, A., Tincani, M., & Frost, L. (2004). Multiply controlled verbal operants: An analysis and extension to the picture exchange communication system. *Behavior Analyst*, 27, 247–261. <https://doi.org/10.1007/BF03393184>.
- Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism*, 2, 325–344. <https://doi.org/10.1177/1362361398024002>.
- Casenhiser, D. M., Binns, A., McGill, F., Morderer, O., & Shanker, S. G. (2015). Measuring and supporting language function for children with autism: Evidence from a randomized control trial of a social-interaction-based therapy. *Journal of Autism and Developmental Disorders*, 45, 846–857. <https://doi.org/10.1007/s10803-014-2242-3>.
- Charlop, M. H., & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied*

- Behavior Analysis*, 22, 275–285. <https://doi.org/10.1901/jaba.1989.22-275>.
- Charlop-Christy, M. H., Le, L., & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disorders*, 30, 537–552. <https://doi.org/10.1023/A:1005635326276>.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provisions for scale disagreement or partial credit. *Psychological Bulletin*, 70, 313–320. <https://doi.org/10.1037/h0026256>.
- Eikeseth, S., & Smith, D. P. (2013). An analysis of verbal stimulus control in intraverbal behavior: Implications for practice and applied research. *The Analysis of Verbal Behavior*, 29, 125–135. <https://doi.org/10.1007/BF03393130>.
- Gena, A., Couloura, S., & Kymissis, E. (2005). Modifying the affective behavior of preschoolers with autism using in-vivo or video modeling and reinforcement contingencies. *Journal of Autism and Developmental Disorders*, 35, 545–556. <https://doi.org/10.1007/s10803-005-0014-9>.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>.
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays*. New York: Allyn & Bacon.
- Hood, S. A., Luczynski, K. C., & Mitteer, D. R. (2017). Toward meaningful outcomes in teaching conversation and greeting skills with individuals with autism spectrum disorder. *Journal of Applied Behavior Analysis*, 50, 459–486. <https://doi.org/10.1002/jaba.388>.
- Ingvarsson, E. T., & Hollobaugh, T. (2010). Acquisition of intraverbal behavior: Teaching children with autism to mand for questions to answers. *Journal of Applied Behavior Analysis*, 43, 1–17. <https://doi.org/10.1901/jaba.2010.43-1>.
- Ishizuka, Y., & Yamamoto, J. (2016). Contingent imitation increases verbal interaction in children with autism spectrum disorders. *Autism*, 20, 1011–1020. <https://doi.org/10.1177/1362361315622856>.
- Jones, C. D., & Schwartz, I. S. (2009). When asking questions is not enough: an observational study of social communication differences in high functioning children with autism. *Journal of Autism and Developmental Disorders*, 39, 432–443. <https://doi.org/10.1007/s10803-008-0642-y>.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Mundy, P., Sigman, M., & Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology*, 6, 389–401. <https://doi.org/10.1017/S0954579400006003>.
- Mundy, P., & Stella, J. (2000). Joint attention, orienting, and nonverbal communication in autism. In A. M. Wetherby & B. M. Prizant (Eds.), *Autism spectrum disorders: A transactional developmental perspective* (pp. 55–77). Baltimore: Brookes.
- Nadig, A., Lee, I., Singh, L., Bosshart, K., & Ozonoff, S. (2010). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia*, 48, 2730–2739. <https://doi.org/10.1016/j.neuropsychologia.2010.05.020>.
- Peters, L. C., & Thompson, R. H. (2015). Teaching children with autism to respond to conversation partners' interest. *Journal of Applied Behavior Analysis*, 48, 544–562. <https://doi.org/10.1002/jaba.235>.
- Schopler, E., Reichler, R. J., & Renner, B. R. (1986). *The childhood autism rating scale (CARS): For diagnostic screening and classification of autism*. New York: Irvington.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton.
- Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of Child Language*, 41, 275–304. <https://doi.org/10.1017/S0305000912000682>.
- Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior*, 27, 23–43. <https://doi.org/10.1007/BF03393090>.
- Tager-Flusberg, H., Calkins, S., Nolin, T., Baumberger, T., Anderson, M., & Chadwick-Dias, A. (1990). A longitudinal study of language acquisition in autistic and down syndrome children. *Journal of Autism and Developmental Disorders*, 20, 1–21. <https://doi.org/10.1007/BF02206853>.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale: Lawrence Erlbaum.
- Ueno, K., Nagoshi, N., & Konuki, S. (2008). *PVT-R Kaiga goi hattatsu kensa [Picture vocabulary test-revised manual]*. Tokyo: Nihon Bunka Kagakusha.
- Vedora, J., Meunier, L., & Mackay, H. (2009). Teaching intraverbal behavior to children with autism: A comparison of textual and echoic prompts. *The Analysis of Verbal Behavior*, 25, 79–86. <https://doi.org/10.1007/BF03393072>.
- Wechsler, D. (1998). *Wechsler intelligence scales for students* (3rd ed.). trans: Azuma, H., Ueno, K., Fujita, K., Maekawa, H., Ishikuma, T., Sato, H.. San Antonio: The Psychological Corporation. (Original work published 1991).
- Wechsler, D. (2010). *Wechsler intelligence scales for children* (4th ed.). trans: Ueno, K., Fujita, K., Maekawa, H., Ishikuma, T., Dairoku, K., Matsuda, O.. San Antonio: Psychological Corporation. (Original work published 2003).