

Effects of Collateral Response Requirements and Exemplar Training on Listener Training Outcomes in Children

Anna Ingeborg Petursdottir · Tracy L. Lepper · Sean P. Peterson

Published online: 9 May 2014
© Association of Behavior Analysis International 2014

Abstract We evaluated the effects of collateral response requirements during listener training on the emergence of vocal foreign-language tacts and intraverbals among 4- and 5-year-old children. In Experiment 1, participants were first exposed to auditory-visual match-to-sample training without collateral response requirements. Four participants did not perform to criterion in probes for derived vocal responses, and were exposed to a two-phase intervention that involved adding echoic and native-language tact requirements to match-to-sample trials. Performance did not improve as a result of the intervention. However, all participants passed tact probes after receiving direct tact and intraverbal training with a subset of the stimuli, and two of four participants also passed the intraverbal probes. Experiment 2 addressed potential limitations of Experiment 1 with three additional participants, but collateral response requirements still failed to affect the emergence of tacts and intraverbals.

Keywords Listener training · Intraverbals · Naming · Tacts · Children

Auditory-visual match-to-sample (MTS) training is commonly used to teach relations between words and their referents; for example, when teaching children with developmental disabilities. A dictated name is presented as a sample stimulus at the beginning of each MTS trial, and the learner's task is to select the corresponding object or a picture from an array of comparison stimuli. Correct responses are differentially reinforced, and incorrect responses may be followed by prompting or error

correction procedures. In applied contexts, auditory-visual MTS training is often considered to be a form of listener training (Greer and Ross 2008; Sundberg 2008), because its ultimate goal is usually to establish control by verbal stimuli over responses that are not verbal “in any special sense” (Skinner 1957, p. 2), such as orienting toward, picking up, or retrieving objects. In addition to its use in early vocabulary instruction, listener training is sometimes a component of instruction in more advanced language and academic skills for learners with and without disabilities (e.g., Joyce and Joyce 1993; Lynch and Cuvo 1995; Melchiori et al. 2000).

A practical limitation of listener training is that it may fail to generate relevant vocal behavior. For example, a child who can point to several colors given their dictated names may be unable to name vocally the same colors upon seeing them. In addition, the child may fail to answer such questions as “What color is grass?” or “Name something that's yellow”, in spite of already being able to name grass and a variety of yellow objects. In Skinner's (1957) terms, the child fails to emit appropriate vocal *tacts* (verbal responses controlled by antecedent nonverbal stimuli) and vocal *intraverbals* (verbal responses controlled by antecedent verbal stimuli, the sound patterns of which differ from those produced by the responses). Failures of listener training to generate vocal repertoires under appropriate stimulus control have been documented in numerous studies with individuals diagnosed with autism and other developmental disabilities (e.g., Lee 1981; Sidman et al. 1986; Wynn and Smith 2003), as well as in typically developing children of preschool- and early school-age (e.g., Connell and McReynolds 1981; Horne et al. 2004; Petursdottir et al. 2008a, b; Petursdottir and Haflidadóttir 2009). However, some of these studies have also included participants who passed tests that required vocal responding. It may be speculated that those participants had prerequisite skills or pre-experimental histories that permitted them to derive greater benefits from the specific listener training

A. I. Petursdottir (✉) · T. L. Lepper · S. P. Peterson
Department of Psychology, Texas Christian University, TCU Box
298920, Fort Worth, TX 76129, USA
e-mail: a.petursdottir@tcu.edu

S. P. Peterson
University of Nebraska Medical Center, Omaha, NE, USA

procedures that were employed in each case. A potentially important avenue of applied investigation involves identifying ways to enhance the effects of listener training on vocal responding, either by building the prerequisite histories, or by modifying the training procedures so that they better match the learner's existing skill set.

Recently, a number of investigators have examined the effects of multiple-exemplar training histories on the emergence of vocal tacts following the training of listener relations or other relations that do not involve vocal responding (Greer et al. 2005, 2007; Luciano et al. 2007; Rosales et al. 2011). Participants in these studies, who have included both typically developing children and children with developmental disabilities, have been exposed to numerous instances of tact and listener training involving the same stimuli (for example, participants might learn to tact a hat, shoe, shirt, and sock, and simultaneously learn to respond as listeners to the names of these stimuli). The researchers have then evaluated the effects of this instructional history on tact emergence following listener training with new stimuli. Although experimental control has been limited in some cases, the results have been promising in that multiple-exemplar training histories have appeared to remediate prior failures of vocal tacts to emerge from listener training alone. These outcomes are consistent with Relational Frame Theory (RFT; Hayes et al. 2001), which proposes histories of training with multiple exemplars as the process by which individuals acquire relational repertoires under appropriate contextual control. However, they are not necessarily incompatible with other accounts of how such repertoires may arise (e.g., Horne and Lowe 1996), nor do they rule out that other types of interventions may produce similar outcomes.

According to Horne and Lowe's (1996) *naming* account, the emergence of vocal tacts following listener training is a product of unsolicited echoic responses to dictated names during training. Horne and Lowe conceptualized naming as a higher-order behavioral relation that involves co-occurrence of tacts, listener behavior, and echoic responses as a result of encountering a stimulus that might evoke one of these relations. Once a child has acquired a repertoire of naming, echoic responses tend to occur collaterally with a child's listener behavior, either at the overt or the covert level. The occurrence of these echoic responses may then permit the visual stimulus selected in a listener training trial to acquire control over the relevant vocal response, resulting in the acquisition of an apparently untrained tact. Horne and Lowe's analysis suggests that one way to enhance the effects of listener training on tact acquisition may be to require an echoic response to the dictated name in each listener training trial. We are aware of two published studies that have demonstrated such an effect. First, Ezell and Goldstein (1989) taught two children with intellectual disabilities to select items from an array and place them in particular locations (e.g., "Put the comb on the chair"). Two

conditions were compared in an alternating-treatments design, one in which the participants were required to echo the dictated stimulus in each trial, and a control condition in which no echoic responses were required. Both participants performed better in subsequent probes for vocal tacts of the experimenter's behavior (e.g., the experimenter placing a comb on a chair) in the echoic condition than in the control condition. Second, Hawkins et al. (2009) reported positive effects of adding an echoic response requirement to the listener trial component of multiple-exemplar training with children diagnosed with autism. However, the effects of the echoic requirement in the absence of multiple-exemplar training were not evaluated.

Given the dearth of published data, we sought to examine the effects of echoic and other collateral response requirements during listener training trials on typically developing children's vocal responding. The evaluation was conducted in the context of teaching foreign-language nouns to preschool- and kindergarten-age children who had already acquired native-language tacts and listener behavior with respect to the visual referents. Listener training in a similar context has typically had limited effects on the emergence of vocal tacts and intraverbals in previous studies with typically developing children of the same age (Petursdottir and Hafliadóttir 2009; Petursdottir et al. 2008b).

Experiment 1

Two collateral response requirements were investigated in Experiment 1. First, we evaluated a requirement to make an echoic response to the foreign dictated name presented in each listener trial. Second, we evaluated the effects of a requirement to follow this foreign-language echoic response with a vocal native-language tact of the visual stimulus selected in that trial. Based on Horne and Lowe's (1996, p. 209) analysis of intraverbal naming, we assessed whether this requirement to emit the foreign- and native-language names of an item in close succession might suffice to establish bidirectional intraverbal relations between the two. Thus, the dependent variables were emergent foreign-language tacts (saying a foreign name given a visual stimulus), native-foreign intraverbals (saying the foreign name of an object given its native-language name, in the absence of a visual stimulus), and foreign-native intraverbals (saying the native-language name of an object given its foreign name, in the absence of a visual stimulus). Tacts and intraverbals were probed in baseline and following completion of a listener training condition in which there were no collateral response requirements. If they did not emerge to criterion, additional listener training was conducted with the same stimuli, during which the two collateral response requirements were added successively to training trials. If tacts and intraverbals still did not emerge to criterion, we

evaluated the effects of training tacts and intraverbals directly with a subset of the stimuli.

Method

Participants

Participants were four children who had no known developmental delays according to parent report. Sophie, Lexi, and Emily were 4 years old and attended a preschool on a full-time basis. Sessions were conducted in the corner of a large room that served as a church library and was not in use during the school day. Ashley was 5 years old and her sessions were conducted in a resource room at her elementary school during after-school care hours. All four participants spoke English at home and at school and did not speak other languages fluently; however, Sophie and Lexi were enrolled in a preschool Spanish class that met once a week during most of their participation in the study.

The experimenter met with each participant for 10–20 min each day that the participant was present and willing to participate, usually three to five times per week. During sessions, the participant and the experimenter were seated side by side at a table. A second observer was sometimes present and seated away from the table. The participant had a token board that required either 18 or 36 tokens to complete. When the participant had filled her token board, the session ended and the participant was given 5-min access to a “fun box” that contained a variety of toys and games (rotated on a regular basis) and occasionally snacks or small items that the participant could take home.

Stimuli

Visual stimuli consisted of 5 cm by 5 cm color photographs obtained from the Picture This© CD-ROM, printed on a white background and framed with black borders. Each participant received instruction with six stimuli and their Japanese names (see Table 1). The visual stimuli were printed on 21.6 cm by 27.9 cm sheets of paper (stimulus sheets) that were inserted into plastic sheet protectors contained in a three-ring binder. The binder contained six stimulus sheets separated by blank sheets. All six stimuli were printed in two horizontally aligned rows on each stimulus sheet, and the location of stimuli on the sheet varied across trials, such that each stimulus appeared once in each of the six locations. In all trials that required visual stimuli, a stimulus sheet was presented by lifting the blank sheet that covered it.

Pre-Experimental Procedures

Native-Language Tact Probes To verify that the participant could tact each visual stimulus in English, the experimenter

presented a trial sheet, pointed to one of the stimuli and asked, “What is this?” Each of the six stimuli was probed once, and the vocalization of any conventional English name for the stimulus resulted in praise and the delivery of a token. The participant’s vocal response determined which English word would later be presented in native-foreign intraverbal probes if more than one acceptable possibility existed. For example, if a participant said “puppy” when presented with a picture of a dog, “puppy” was used in all native-foreign intraverbal probes for that participant. Had a participant failed to vocalize any conventional English name for a particular stimulus, that stimulus would have been replaced with another stimulus that the participant could tact in English. However, this never happened.

Native-Language Listener Pretraining Following the native-language tact probe, brief training was conducted to ensure that the participant could scan the array of stimuli on the trial sheet and select an appropriate stimulus in response to its dictated English name. Before training began, the experimenter said, “I am going to say some words, and I want you to point to the picture if you can.” The experimenter then presented the English names of the stimuli one by one (e.g., “train”), each followed by the presentation of a trial sheet. The English names presented matched the participant’s responses in the preceding native-language tact probe (e.g., “puppy” for a dog, if the participant had said “puppy” in the tact probe). A correct response, defined as touching the correct stimulus within 5 s without first touching another stimulus, resulted in praise and the delivery of a token. An incorrect response was followed by a pointing prompt and repetition of the trial. Training continued until the participant responded with 100 % accuracy in three consecutive six-trial blocks, in which each block contained one presentation of each stimulus and presentation order varied across blocks. All participants met this criterion in the minimum of three trial blocks.

Foreign-Language Echoic Pretraining Echoic pretraining was conducted to ensure that the participant could echo the target Japanese names. Before training began, the experimenter said, “I am going to tell you some Japanese words. Japanese is a foreign language. I am going to say the words, and then I want you to say them too”. The experimenter vocally presented the Japanese names one by one (e.g., “kisha”) and waited for a response. A correct response was defined as an exact vocal match or a close approximation (e.g., substitution of one consonant; if a participant made consistent substitutions, they were subsequently accepted as correct responses for that participant during the experiment), and was followed by praise and the delivery of a token. If no response or an incorrect response was made, the experimenter presented the name again more slowly. Had a participant been unable to respond correctly at this time, the stimulus would have been replaced

Table 1 English and Japanese Names for Visual Stimuli in Experiment 1

Participant	English Names	Japanese Names
Sophie, Emily, Lexi	Cow, dog (puppy, doggie), flower, train, carrot, apple	Ushi, inu, hana, kisha, ninjin, ringo
Ashley	Bird, monkey, cat, strawberry, lettuce, potato	Tori, saru, neko, ichigo, jagaimo, retasu

with another Japanese name that the participant could echo, but this was never necessary. Training continued until the participant responded with 100 % accuracy in three consecutive six-trial blocks, in which each block contained one presentation of each stimulus. All participants met the criterion in the minimum of three blocks.

Dependent Variables and Data Collection

The primary dependent variables were emergent verbal relations, defined as correct responses in probes for foreign tacts and two types of intraverbals; native-foreign and foreign-native. These probes were accompanied by nonreinforced probes for the six listener relations that were targeted in all training conditions. Table 2 shows the antecedent stimuli presented in the four types of probe trials, along with target response definitions. During probe sessions, the experimenter recorded correct and incorrect responses on a data sheet. In all probe trials that required a vocal response (tact, foreign-native, and native-foreign trials), a response was scored as correct if the participant vocalized the target name within 5 s of the experimenter's instruction. When the target response was an English name (foreign-native trials), an incorrect response was scored if the participant vocalized a different English name or did not vocalize any English name within 5 s. When the target response was a Japanese name (tact and native-foreign trials), an incorrect response was scored if the participant vocalized a different Japanese name or did not vocalize any Japanese name within 5 s. If the participant vocalized two or more names in the specified language in the same trial (e.g., two Japanese names in a tact trial), only the first was scored as correct or incorrect. In listener probe trials, a response was scored correct if the participant touched the visual stimulus corresponding to the Japanese name presented by the experimenter within 5 s of the instruction, without first touching another stimulus. An incorrect response was scored if the participant touched a different stimulus or did not touch a stimulus within 5 s.

Secondary dependent measures included the acquisition (trials to criterion) of foreign-language listener responses, the occurrence of echoic responses during baseline and training trials, and the occurrence of native-language tacts during baseline and training trials. The experimenter recorded listener responses on a data sheet during all training sessions, and correct and incorrect responses were defined in the same manner as in probe trials. Data on the occurrence of echoic

responses and tacts were recorded on a data sheet from video. The observer scored an echoic response if the participant repeated the Japanese name presented by the experimenter without any prompting. A native-language tact was scored if the participant vocalized the English name of the visual stimulus that he or she selected, without prompting.

Interobserver Agreement

An independent observer collected data during at least 30 % of each participant's probes, either live or from video. An agreement was scored for each trial in which both observers recorded a correct response or both recorded an incorrect response; a disagreement was scored if the observers' records differed. Interobserver agreement (IOA) was calculated for each probe session by dividing the number of agreements by the sum of agreements and disagreements and converting the ratio to a percentage. Mean IOA was 100 % for Sophie, Emily, and Ashley, and 97.2 % (range 91.6 % to 100 %) for Lexi.

IOA was also assessed for 60 % of each participant's baseline and training trial blocks. IOA for each trial block was calculated in the same manner as for probes. Mean IOA was 95.4 % (range 66.7 % to 100 %) for Sophie, 99.5 % (range 66.7 % to 100 %) for Lexi, 99.5 % (range 66.7 % to 100 %) for Emily, and 99.8 % (range 83.3 % to 100 %) for Ashley. Finally, IOA on echoic responses and native-language tacts was assessed for at least 60 % of each participant's baseline and training trial blocks. Occurrence IOA for each trial block was calculated by dividing the number of trials in which both observers recorded a response by the number of trials in which at least one observer recorded a response and converting this ratio to a percentage. Occurrence IOA on echoic responses was 94.6 % (range 50 % to 100 %) for Sophie, 77.3 % (range 0 % to 100 %) for Lexi, 95.1 % (range 0 % to 100 %) for Emily, and 94.6 % (range 0 % to 100 %) for Ashley. Occurrence IOA on native-language tacts was 95.3 % (range 0 % to 100 %) for Sophie, 73.1 % (range 0 % to 100 %) for Lexi, 97.5 % (range 60.0 % to 100 %) for Emily, and 98.3 % (range 83.3 % to 100 %) for Ashley. Zero percent IOA occurred in some sessions during SLT when one observer recorded one echoic or tact response, and the other recorded zero.

Procedure

Overview Following pre-experimental procedures, a listener baseline was conducted, followed by baseline probes of the

Table 2 Antecedent Stimuli and Correct Response Definitions for Probe Trials

Probe	Antecedent stimuli	Correct response
Listener relation	Experimenter vocalizes Japanese name and uncovers stimulus sheet	Touch visual stimulus within 5 s
Foreign tact	Experimenter points to visual stimulus and asks, “What is this in Japanese?”	Vocalize Japanese name within 5 s
Native-foreign intraverbal	Experimenter asks, “What is Japanese for [English word]?”; no visual stimuli are in sight.	Vocalize Japanese name within 5 s
Foreign-native intraverbal	Experimenter asks, “What is [Japanese word]?”; no visual stimuli are in sight.	Vocalize English name within 5 s

relations shown in Table 2. Participants then underwent the following sequence of training conditions: (a) standard listener training, (b) Collateral Response Training 1 (CRT-1), (c) Collateral Response Training 2 (CRT-2), and (d) exemplar training, which consisted of up to four phases. When the training criterion was met in each training condition or in each phase of exemplar training, the probes shown in Table 2 were repeated. The participant proceeded to the next training condition or the next phase of exemplar training if, in the probes, she (a) responded correctly in at least five out of six listener probe trials, and (b) responded correctly in fewer than five out of six probe trials for at least one of the three vocal relations. If the participant made fewer than five correct listener responses, indicating that the trained listener repertoire was not maintained under probe conditions, the participant returned to the previous training condition, which continued until the training criterion was met again, followed by repeated probes. In addition, the experimental design required some participants to return to the previous training condition for additional training in spite of having met the criterion of five correct listener responses. If in any probe session, the participant made at least five correct listener responses, and also responded correctly on at least five foreign tact, five foreign-native, and five native-foreign trials, this concluded her participation.

Experimental Design Probe data were used to evaluate the effects of CRT-1, CRT-2 and exemplar training on emergent vocal relations. A nonconcurrent multiple-probe design across participants was used to evaluate the effects of exemplar training, while controlling for the amount of previous instruction. Thus, the introduction of exemplar training was staggered across post-CRT baselines that differed in length. It is important to note that the extended post-CRT baselines for Emily, Ashley, and Lexi did not consist of simply of repeating the probes multiple times following a single training phase. Rather, each post-CRT probe was preceded by additional CRT to criterion in order to control for additional exposure to training.

Probes A probe session consisted of 24 trials that included six listener, six foreign tact, six foreign-native, and six native-foreign trials; thus, there was one trial per stimulus for each

type of relation. The sequence of trials varied across probe sessions, and within each session, the sequence was quasi-randomized such that every four trials included one trial for each type of relation, but not necessarily with the same stimuli. Every six to ten trials, the participant received a brief break from probe trials, during which the experimenter delivered several tokens that were noncontingent on performance in the session. The antecedent stimuli presented in probe trials are shown in Table 2. Following stimulus presentation, the experimenter waited up to 5 s for a response and then initiated the next trial. In foreign tact trials, if the child responded with the English rather than the Japanese name associated with the visual stimulus, the experimenter prompted a Japanese response by asking, “What is it in Japanese?” with an emphasis on the last word. If the participant vocalized the appropriate Japanese response at that time, it was scored as correct in spite of the prompt. Otherwise, no feedback or other consequences were provided for either correct or incorrect responses. A particular type of relation was considered to have emerged to criterion if the participant responded correctly on five out of six trials for that relation in a single probe session (with six comparison stimuli, the probability of meeting this criterion by chance is <0.0001).

Listener Baseline A baseline of listener responding, as well as echoic and native tact responses on listener trials, was conducted in six-trial blocks in which each block contained one listener trial per stimulus. Presentation order was determined by the experimenter’s data sheet and varied across trial blocks. At the beginning of each session, the experimenter informed the child that “Today we are going to practice some Japanese words. I will say the word and you point to the picture if you can.” Baseline listener trials were identical to probe trials for listener responses (see Table 2), and as in probe trials, the Japanese name was presented without an instruction in each trial. A listener baseline was accidentally omitted for Lexi.

Standard Listener Training Before standard listener training began, the training stimuli were divided into three two-stimulus sets. Standard listener training was then conducted in four steps. In the first step, training was conducted with the first set until the participant responded correctly in a single six-trial block that contained three presentations of each

stimulus. This procedure was repeated with the remaining stimulus sets in the next two steps. In the final step, presentations of all six stimuli were intermixed, as in baseline. This step continued until the participant responded correctly in at least 17 of the 18 trials that comprised three consecutive trial blocks.

At the beginning of each session the experimenter informed the child that “Today we are going to practice our Japanese words. I will say the word and you point to the picture if you can.” Standard listener training trials were identical to baseline trials except for the addition of prompting procedures, differential reinforcement of correct responses, and error correction. In the first trial block of Phases 1, 2, and 3, the experimenter delivered an immediate prompt in each trial by pointing to the correct visual stimulus immediately after saying the Japanese word and uncovering the stimulus sheet. The experimenter praised all prompted responses and delivered a token. In the second trial block, the delay to the pointing prompt was increased to 2 s from the presentation of the stimulus sheet, and the experimenter delivered praise and tokens for both prompted responses and correct responses. From the third trial block on, as well as throughout all of Phase 4, the prompt delay was 5 s, and praise and tokens were contingent on correct, unprompted responses. Following each prompted response, error correction was implemented by repeating the trial with a new stimulus sheet until a correct response was obtained without prompting.

CRT-1 CRT-1 was identical to Phase 4 of standard listener training, with the addition of a requirement for an echoic response in each trial. Before each session, the experimenter told the participant that “Today we’re going to practice our Japanese words. I will say the word, then I want you to say the word too, and then point to the picture.” Following the presentation of the Japanese word, the experimenter waited for an echoic response. If the participant did not make an echoic response within 5 s, the experimenter prompted an echoic response by saying, “Say [Japanese word]”. Immediately following a prompted or an unprompted echoic response, the experimenter uncovered the trial sheet and waited up to 5 s for a listener response. As in standard listener training, a correct listener response resulted in praise and the delivery of a token, whereas an incorrect listener response was followed by a pointing prompt and the trial was repeated. The praise and tokens were contingent only on correct listener responses, and not on the occurrence of unprompted echoic responses. CRT-1 continued until three consecutive trial blocks were completed in which the participant (a) made a correct listener response in at least 17 out of 18 trials, and (b) made an unprompted echoic response in at least 17 out of 18 trials.

CRT-2 CRT-2 was identical to CRT-1 with the addition of a native tact requirement in each trial. Before each session, the

experimenter told the participant that “Today we’re going to practice our Japanese words like we did before. I will say the word, then you will say the word, and when you point to the picture, I want you also to tell me what you usually call it.” Trials proceeded as in CRT-1, with the exception that the participant was now required to vocalize a native-language tact of the selected visual stimulus within 5 s of touching it. If the child failed to vocalize a native-language tact or the native-language tact was incorrect, the experimenter prompted a native-language tact by saying, “Point again and tell me what you usually call it.” If a correct native-language tact still did not occur within 5 s of touching the visual stimulus, the experimenter modeled the response by pointing to the visual stimulus and simultaneously saying its English name. Following a prompted tact, the trial was repeated, regardless of whether the listener response was correct or incorrect. CRT-2 continued until three consecutive trial blocks were completed in which the participant (a) made a correct listener response in at least 17 of the 18 trials, (b) made an unprompted echoic response in at least 17 out of 18 trials, and (c) emitted an unprompted native-language tact in at least 17 out of the 18 trials.

Exemplar Training Exemplar training consisted of directly training foreign tacts, native-foreign intraverbals and foreign-native intraverbals with consecutive subsets of the stimuli, followed by training with all six stimuli if necessary. For Phases 1 through 3, the six stimuli were divided into three two-stimulus sets (not necessarily the same sets as those used in the first three steps of standard listener training), and one set was employed in each phase. In Phases 1 through 3, each trial block consisted of two standard listener, two foreign tact, two native-foreign and two foreign-native trials. Stimulus presentation was identical to probe trials (see Table 2). Correct responses were followed by praise and a token. Following an incorrect response, the experimenter prompted a correct response and repeated the trial until the participant made a correct response without prompting. Training continued until the participant made at least seven correct responses in each of three consecutive trial blocks. In Phase 4, training was conducted with all six stimuli, and each trial block consisted of six standard listener, six foreign tact, six native-foreign, and six native-foreign trials. Phase 4 training continued until the participant made at least five correct responses for each of the four types of relations in three consecutive trial blocks. Probes were conducted following each phase, and participants proceeded to the next phase only if they did not pass probes for one or more of the three vocal relations (i.e., foreign tacts, native-foreign intraverbals, or foreign-native intraverbals). Before the probes that followed Phases 1, 2, and 3, all six listener relations were probed once without feedback. If the participant made an error in this block of listener trials, up to two additional blocks were conducted, and if the participant

made one or more errors in all three blocks, the participant returned to standard listener training, which was completed to criterion before proceeding to the probes.

Results

Training Data

Sophie, Emily, and Ashley performed at chance level in the listener baseline (data available from first author upon request). Table 3 shows each participant's trials to criterion in standard listener training, CRT-1, and CRT-2. The number of trials required to complete standard listener training ranged from 210 to 636, or 35 to 106 trial blocks. Sophie initially met the standard listener training criterion after 294 trials, but in the subsequent probe, she did not meet the criterion of five correct listener responses. As a result, five additional blocks of training trials were conducted until the criterion was met again, for a total of 324 trials. Similarly, Lexi first met the criterion after 606 trials, and also required five additional blocks of training to meet the criterion again after failing the listener portion of the probe.

The number of trials to criterion in CRT-1 ranged from 36 to 132, and from 18 to 90 in CRT-2. In CRT-1, Lexi initially met the criterion after 18 trials, but required 19 additional trial blocks (for a total of 132 trials) to meet the criterion again after failing the listener portion of the subsequent probe. In line with the requirements of the multiple-probe design, Lexi and Emily each received an additional round of CRT-2 to criterion after initially completing training and passing the listener portion of the subsequent probe. Ashley received one additional round of CRT-1 and two additional rounds of CRT-2. These overtraining phases were completed in 18 to 24 trials, except for Ashley's CRT-1 overtraining, which took 156 trials to complete.

Finally, all participants were exposed to Phase 1 of exemplar training, and Emily and Ashley were exposed to additional phases of exemplar training. Trials to criterion in each phase are shown in Table 3. Following Phase 1, Lexi failed to maintain the six listener relations. As a result, she returned to standard listener training, which took 384 trials to complete to criterion, before proceeding to the probe for tact, native-foreign intraverbal and foreign-native intraverbal relations. The other participants maintained the six listener relations throughout exemplar training, and did not require a return to standard listener training.

Table 4 shows the percentage of trials with unprompted echoic responses and native-language tacts in the listener baseline, standard listener training, CRT-1, and CRT-2. In baseline, Sophie tacted the visual stimuli she selected in English in 26.2 % of all trials, but did not make any echoic responses. Emily, by contrast, echoed most of the Japanese names but did not tact the visual stimuli. Ashley did not emit

any echoic responses or native-language tacts. During standard listener training, some echoic responses and native-language tacts were observed for all participants. These responses generally occurred at very low levels, with the exception that Lexi and Emily emitted echoic responses in approximately a fifth of their trials. For Emily, echoic responses occurred at high rates early in listener training and fell to zero after the first few blocks of intermixed trials (details available from first author upon request). Lexi, by contrast, emitted echoic responses at similar levels throughout training. During CRT-1, there was a large increase in unprompted echoic responses, but not in native-language tacts, for all participants. During CRT-2, both echoic responses and native-language tacts occurred in a large percentage of all trials.

Probe Data

Figure 1 shows the probe performance of all participants; grey bars represent trained listener responses and markers represent untrained vocal relations. After baseline, the figure shows only data from probe sessions in which the participant met the criterion of five out of six correct listener responses (i.e., we have omitted data from one probe for Sophie and three probes for Lexi in which the listener relations were not maintained, suggesting a need for continued training). In baseline, all participants responded around chance level in listener probes, and emitted few or no correct responses in foreign tact, native-foreign, and foreign-native trials. Emily responded correctly in two foreign tact trials; this was due to her responses in most trials alternating between two of the Japanese names that she had been exposed to in the listener baseline and continued to be exposed to in listener and foreign-native intraverbal probe trials. Following standard listener training, all participants responded with increased accuracy when probed for the three vocal relations. Sophie and Ashley met the criterion for emergent foreign-native intraverbals, but did not meet the criterion for foreign tacts or native-foreign intraverbals. Lexi and Emily did not meet criterion for any of the relations. Following CRT-1, all participants' correct responses increased slightly in foreign tact or native-foreign intraverbal trials, but no participant met the criterion for either relation. Lexi met the criterion for foreign-native intraverbals, but Emily still did not. Ashley received an additional round of CRT-1 to criterion after meeting the training and listener probe criteria the first time, but her performance did not improve to criterion. Following CRT-2, Emily met the criterion for emergent foreign-native relations, but there were no increases in any participants' accuracy in foreign tact or native-foreign trials. Lexi and Emily received one additional round, and Ashley received two additional rounds of CRT-2 after the first time they met the training and listener probe criteria. None of these participants' performance improved with repeated CRT-

Table 3 Trials to Criterion during Training in Experiment 1

	Standard Listener Training	CRT-1	CRT-1 over-training	CRT-2	CRT-2 over-training	Exemplar Training			
						Phase 1	Phase 2	Phase 3	Phase 4
Sophie	324	36	-	90	-	104	-	-	-
Lexi	636	132	-	60	24	448 ¹	-	-	-
Emily	216	66	-	60	18	24	32	32	144
Ashley	210	36	156	18	18, 18	64	24	-	-

¹ Followed by 384 trials of standard listener training

2 training and testing. Finally, all participants received exemplar training. All four participants met the criterion for foreign tacts after receiving exemplar training with only two of the six stimuli (i.e., after completing Phase 1), and Sophie and Lexi also met the criteria for native-foreign intraverbals. Ashley and Emily received exemplar training with two additional stimuli in Phase 2, following which Ashley met the criterion for native-foreign intraverbals, but Emily did not. Emily went on to receive exemplar training with the remaining two stimuli in Phase 3, but did not meet the native-foreign criterion until she had also received exemplar training with all six stimuli simultaneously in Phase 4.

Discussion

In summary, the introduction of collateral response requirements into listener training did not substantially enhance its effects on the emergence of tacts and intraverbals. The initial completion of standard listener training without collateral response requirements was followed by increases in correct vocal responses from baseline, and two of four participants achieved criterion performance with foreign-native relations. However, consistent with prior research (Petursdottir and Hafliadóttir 2009; Petursdottir et al. 2008b), performance in foreign-tact and native-foreign intraverbal probes remained below criterion for all participants. The addition of collateral response requirements to the listener training protocol in no case sufficed to bring about criterion performance with all relations. Even though some of the participants were exposed to repeated rounds of collateral response training, no increases

to criterion were observed in correct responses in foreign tact and native-foreign probes following either CRT-1 or CRT-2. Only after tacts and intraverbals were trained directly with at least a subset of the stimuli was criterion-level performance observed in probes for these relations.

One possible reason for the lack of an effect of CRT is that for all participants except Ashley, the CRT phases were very brief compared to standard listener training. Because the participants had already acquired the listener repertoire during standard listener training, little additional training was needed during CRT-1 and CRT-2 before the collateral responses occurred reliably along with the listener responses. Although some participants were exposed to extended CRT as part of the multiple-probe design, the total amount of exposure ranged from only 36 to 192 trials in CRT-1, and from 54 to 90 trials in CRT-2. By contrast, standard listener training ranged from 210 to 636 trials. An effect might have been observed if CRT had continued longer or if collateral responses had been required from the beginning of listener training. Another possible limitation is that a minimal number of probe trials were conducted in each condition, which may have resulted in variability unrelated to the training conditions (e.g., see Lexi's data following the two iterations of CRT-2). These limitations were addressed in Experiment 2.

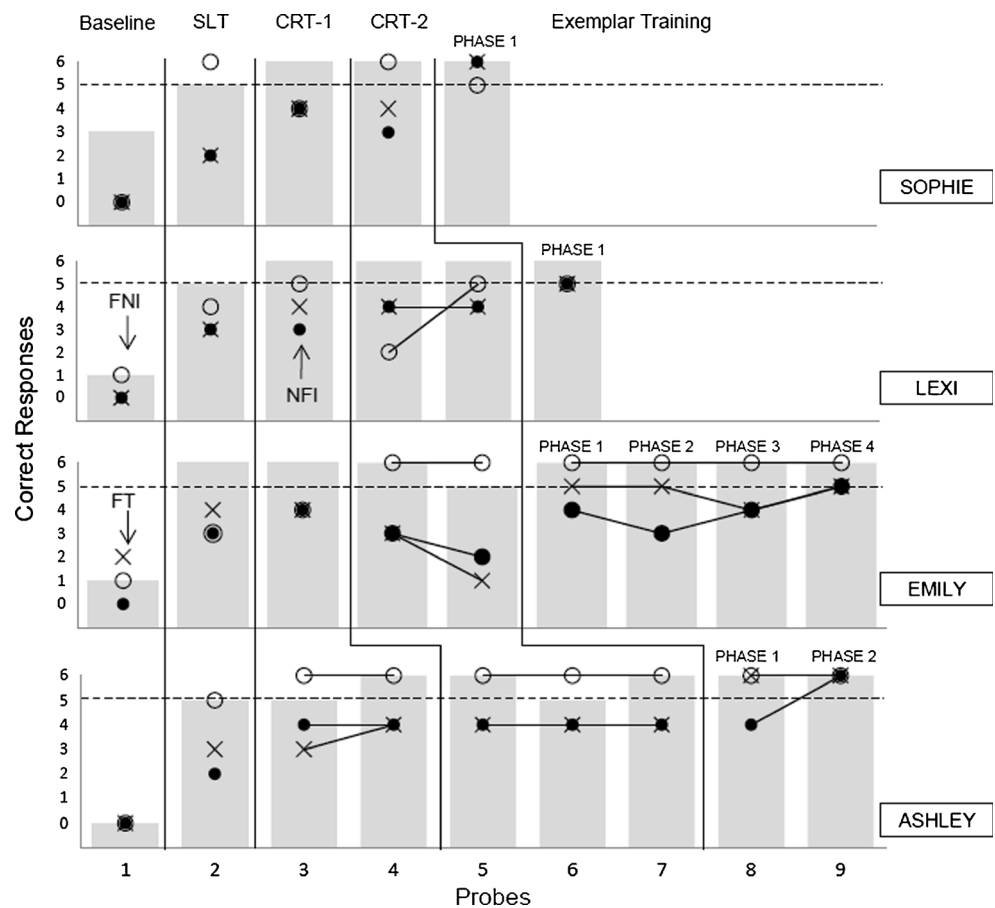
Experiment 2

In Experiment 2, participants received listener training with two three-stimulus sets concurrently. One stimulus set was

Table 4 Percentage of Trials with Unprompted Echoic Responses and Native-Language Tacts in Experiment 1

	Listener baseline		Standard listener training		CRT-1		CRT-2	
	Echoic	Native Tact	Echoic	Native Tact	Echoic	Native Tact	Echoic	Native Tact
Sophie	0.0 %	26.2 %	6.5 %	4.9 %	88.9 %	2.8 %	98.9 %	77.8 %
Lexi	-	-	19.8 %	2.0 %	95.5 %	1.3 %	100 %	84.5 %
Emily	94.7 %	0.0 %	21.8 %	2.8 %	100 %	0.0 %	98.7 %	84.6 %
Ashley	0.0 %	0.0 %	1.9 %	1.0 %	99.5 %	0.5 %	100 %	98.1 %

Fig. 1 The performance of participants in Experiment 1 in probes conducted in baseline and following each training condition. Grey bars represent trained listener responses. SLT=Standard Listener Training; CRT= Collateral Response Training; FT=Foreign Tact; FNI=Native-Foreign Intraverbal; FNI= Foreign-Native Intraverbal The dashed lines represent the criterion of five of six correct responses. Data from probes in which the participants made fewer than five correct listener responses are omitted, as they resulted in continued training followed by another post-training probe



assigned to a training condition that was equivalent to CRT-1 in Experiment 1 and the other was assigned to standard listener training. Tacts and intraverbals were probed throughout training. In Experiment 2, we evaluated only the echoic response requirement, and not the native-language tact requirement, because (a) to the extent that increases in probe performance were seen in Experiment 1, they occurred following CRT-1 and not following CRT-2, and (b) the effects of the echoic response requirement have been documented in other studies (Ezell and Goldstein 1989; Hawkins et al. 2009). Thus, the purpose of Experiment 2 was to investigate whether the effects of CRT-1 that have been observed in other studies, but were not observed in Experiment 1, could be captured in a different experimental design.

Method

Participants

Three 4-year-old girls participated. Maia, Lauren, and Jennifer had no known developmental delays according to parent report, spoke English in their homes, attended a preschool on a full-time basis, and were enrolled in a preschool Spanish class that met once a week. The setting was identical to that described for Sophie, Emily, and Lexi in Experiment 1.

Stimuli

In Experiment 2, each participant initially received instruction with six stimuli divided into two three-stimulus subsets (sets 1a and 1b) that were randomly assigned to standard listener training and CRT-1. Maia and Lauren then underwent a second instructional phase with new stimuli (sets 2a and 2b) that were assigned to the same two conditions. Table 5 lists all stimuli used for each participant.

The visual stimuli were similar to those used in Experiment 1. In each training condition, each trial sheet contained the three instructional stimuli assigned to that condition, and three additional visual stimuli that were not targeted in training and served only to equate the number of comparisons per trial to those in Experiment 1.

Data Collection and Interobserver Agreement

Response definitions and coding were identical to Experiment 1. IOA was assessed for at least 30 % of all probes and at least 30 % of all baseline and training sessions, in a manner identical to Experiment 1. During probes, mean IOA was 100 % for Maia, 99.4 % (range, 97.2 % to 100 %) for Lauren, and 99.5 % (range, 97.2 % to 100 %) for Jennifer. During baseline and training, mean IOA on listener responses was 99.6 %

Table 5 English and Japanese Names for Visual Stimuli in Experiment 2

Participant	Set 1a Standard listener training		Set 1b CRT-1		Set 2a Standard listener training		Set 2b CRT-1	
	English	Japanese	English	Japanese	English	Japanese	English	Japanese
Maia	glasses	megane	butterfly	cho	cat	neko	bird	tori
	dog	inu	flower	hana	monkey	saru	lettuce	retasu
	apple	ringo	clock	tokei	straw-berry	ichigo	potato	jagaimo
Lauren	bird	tori	carrot	ninjin	butterfly	cho	glasses	megane
	horse	uma	cow	ushi	flower	hana	dog	inu
	egg	tamago	train	kisha	clock	tokei	apple	ringo
Jennifer	bird	tori	cat	neko	-	-	-	-
	lettuce	retasu	monkey	saru				
	potato	jagaimo	straw-berry	ichigo				

(range, 93.3 % to 100 %) for Maia, 100 % for Lauren, and 99.7 % (range, 93.3 % to 100 %) for Jennifer. Mean occurrence IOA for echoic responses was 91.1 % (range, 86.7 % to 100 %) for Maia, 98.7 % (range, 93.3 % to 100 %) for Lauren, and 100 % for Jennifer.

Procedure

Following pre-experimental procedures (identical to Experiment 1), standard listener training and CRT-1 were compared in an adapted alternating-treatments design (Sindelar et al. 1985). A nonconcurrent multiple-baseline design across participants was used to control for possible acquisition outside of the experiment. Maia and Lauren underwent two comparisons of standard listener training and CRT-1 with different stimulus sets, whereas Jennifer underwent only one evaluation due to time constraints.

Throughout baseline and training, sessions alternated across the standard listener training and CRT-1 conditions such that each pair of sessions contained one session in each condition, but their order varied across session pairs. Each session consisted of 15 trials; five trials per each of the three stimuli that were assigned to the condition in effect. Each pair of training sessions was followed by a probe session for foreign tacts, foreign-native intraverbals, and native-foreign intraverbals (the trained listener relations were not probed in this experiment). Each probe session included all six stimuli assigned to both training conditions, and each of the three relations was probed twice per stimulus for a total of 36 trials (18 in each condition).

The mastery criterion for standard listener training and CRT-1 was five out of six correct responses for each of the three stimuli in three consecutive sessions. When this criterion was met in one of the training conditions, training ceased in that condition, and from then on, training sessions in the other condition alternated with probe sessions.

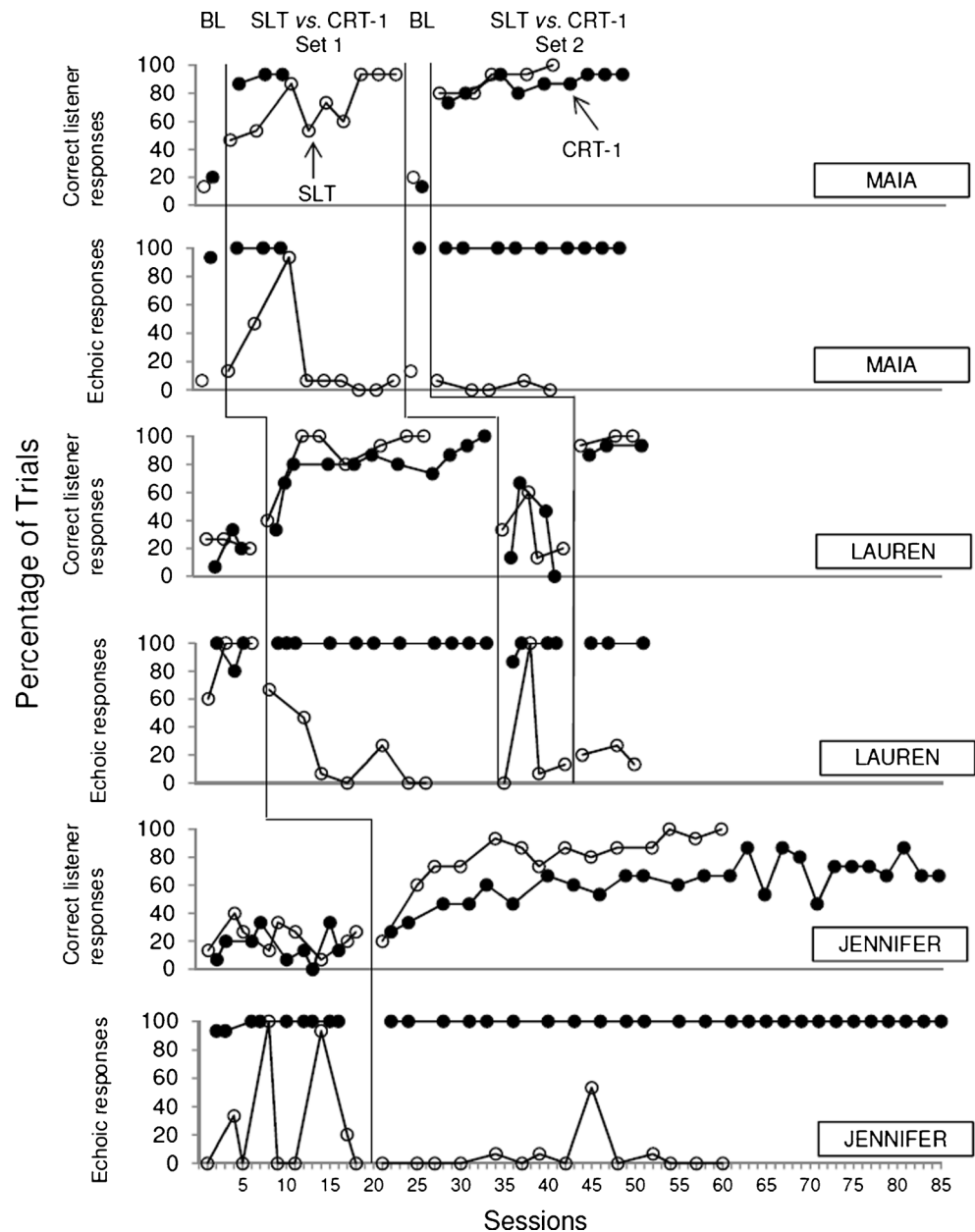
Stimulus presentation, prompts, and consequences employed during probes, baseline, standard listener training, and CRT-1 were identical to Experiment 1, with the exception that the baseline for the CRT-1 subsets included a requirement to make an echoic response in each trial that was implemented in the same manner as during CRT-1 training.

Results and Discussion

Figure 2 shows correct listener and echoic responses for all participants during baseline and training. With stimulus Set 1, Maia required fewer sessions to meet criterion in CRT-1 than in standard listener training, whereas with Set 2, she required fewer sessions in standard listener training. Lauren required fewer sessions in standard listener training than in CRT-1 with Set 1, but an equal number with Set 2. Jennifer met the training criterion only in the standard listener training condition. Thus, there was no evidence that the echoic response requirement included in CRT-1 facilitated acquisition of listener relations. Echoic responses were seen primarily in the CRT-1 condition. Lauren and Jennifer made a number of echoic responses in the baseline of the standard listener condition, but these responses dropped out prior to or during training.

Figure 3 shows the participants' probe performance following the listener baseline and during training until mastery was reached in each training condition. With Set 1, Maia made few correct responses during probes, except that she met the criterion (five of six correct responses) for foreign-native intraverbals with the standard listener training stimuli at the end of standard listener training. With Set 2, Maia's performance in both conditions improved from the first set. Foreign-native intraverbals emerged to criterion with stimuli from both conditions, and tacts with stimuli from the CRT-1 condition. Overall, Maia's terminal performance at the end of training did not differ across training conditions. For Lauren, foreign-native intraverbals emerged to criterion with

Fig. 2 Training data for participants in Experiment 2. BL=Baseline; SLT=Standard Listener Training. The upper panel for each participant shows the percentage of correct listener responses and the lower panel shows the percentage of trials with an echoic response



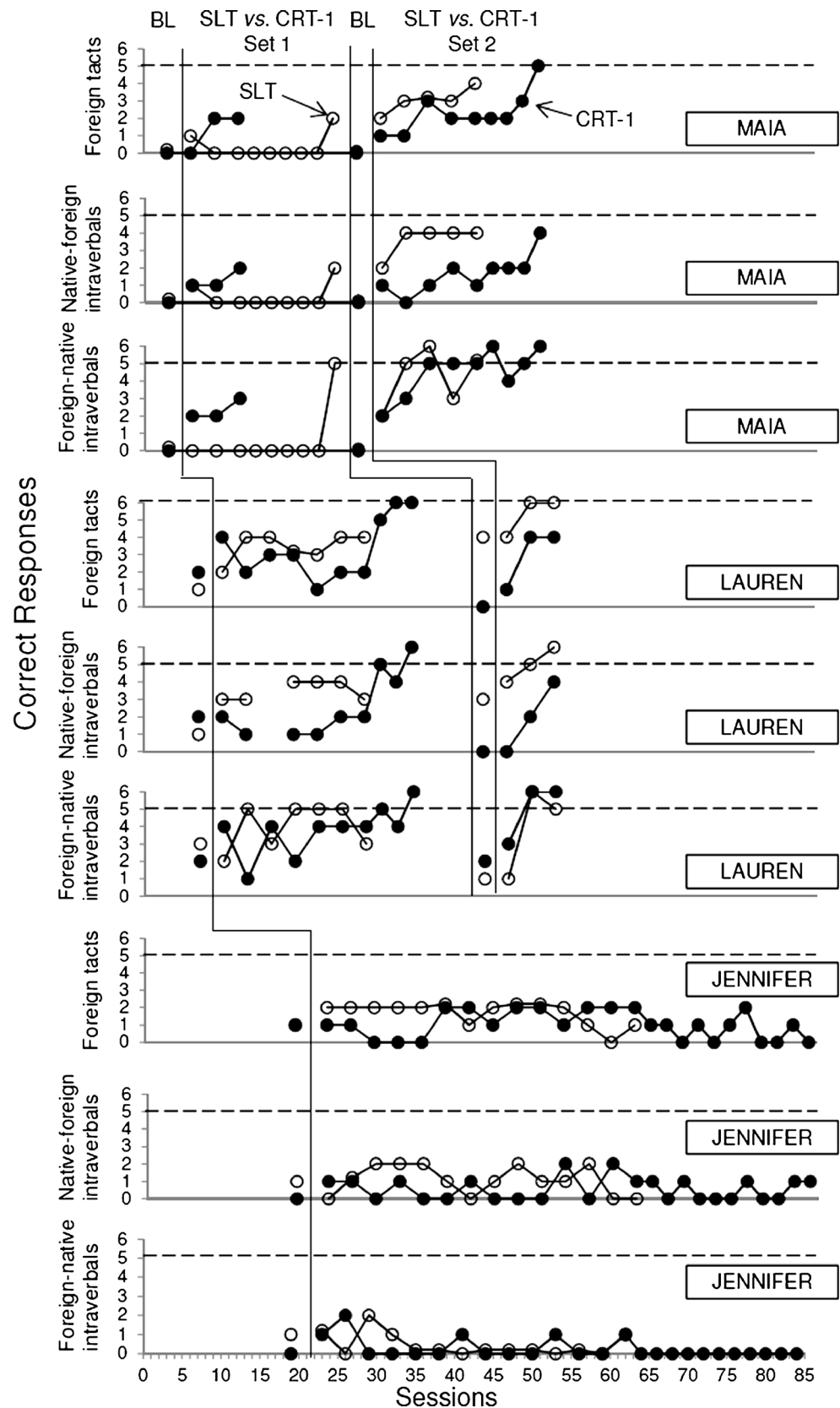
stimuli from both conditions and with both stimulus sets. With Set 1, tacts and native-foreign intraverbals emerged to criterion with the CRT-1 stimuli but not with the standard listener training stimuli; however, the opposite results were seen with Set 2. Jennifer made few correct responses in probes for all relations; there was no difference between conditions and no type of relation emerged to criterion.

Following mastery in one training condition, we continued to include stimuli from that condition in probe sessions until mastery was reached in the other condition, even though training had ceased. However, Fig. 3 includes data only from the first post-mastery probe session, as we were

most interested in the participants' performance at the time of mastery. For Maia and Lauren, probe performance with stimuli from the mastered condition generally increased in accuracy following the first post-mastery probe (data available from first author upon request). However, the post-mastery probe data did not alter our conclusion that there were no systematic differences between conditions for any participant.

In summary, collateral echoic response requirements failed to enhance the effects of listener training on untrained tacts and intraverbals, even when included from the beginning of training. We did not necessarily expect the emergence of intraverbals to be affected by the manipulation in Experiment

Fig. 3 Probe data for participants in Experiment 2. The dashed lines represent the criterion of five of six correct responses. Data on foreign-native intraverbals are missing from Lauren’s third probe during training with Set 1 due to a probe administration error



2, because based on Horne and Lowe’s (1996) analysis, we expected the echoic response to primarily affect the emergence

of foreign tacts. However, tact emergence was not affected to a greater degree than the emergence of intraverbals.

General Discussion

Consistent with previous research, listener training generated some emergent vocal tacts and intraverbals. Foreign-native intraverbals often emerged to criterion (Sohpie and Ashley in Experiment 1; Maia and Lauren in Experiment 2), but foreign tacts and native-foreign intraverbals typically did not (the only exception was Lauren in Experiment 2). The addition of collateral response requirements to the standard listener training protocol failed to increase accuracy in probes for tacts and intraverbals, regardless of whether these requirements were introduced after the successful completion of standard listener training (Experiment 1) or included from the beginning of training (Experiment 2).

The absence of an effect of collateral response training may seem surprising given that the CRT conditions actually included a contingency on emitting the foreign-language vocal response topographies in close contiguity with the presentation of stimuli that might be expected to evoke them as tacts or native-foreign intraverbals later. However, CRT was not designed to guarantee that tacts or intraverbals would be established, as no attempt was made actually to transfer control over the foreign-language responses from the words dictated by the experimenter to the relevant visual stimuli or native-language names. Such attempts were deliberately left out of the CRT protocols, as Horne and Lowe's (1996) analysis implies that they should not be necessary in the natural environment. However, it appeared that control over echoic responses did not transfer automatically to the visual stimuli, or at least not to all of the visual stimuli.

The results appear inconsistent with previously published studies on the effects of echoic responses during listener training on emergent vocal behavior (Ezell and Goldstein 1989; Hawkins et al. 2009). One possible explanation is that the participants in these prior studies were children who had language impairments due to developmental disabilities. Horne et al. (2004) reported an unpublished study by Bell (1999) that demonstrated a similar effect with typically developing children, but the participants in that study were much younger than those in the present study (i.e., 20 to 23 months old). In line with Horne and Lowe's (1996) analysis of naming and the timeline of its acquisition, it may be speculated that the typically developing 4- and 5-year-old participants in the present study (but not participants in previous studies) already had a strong naming repertoire, and thus were already echoing the dictated sample stimuli during training (and possibly making other collateral responses, such as tacting the selected visual stimuli in their native language), but covertly. If that was the case, the requirement to make these responses at the overt level may not have exerted further effects on their probe performance. According to Horne and Lowe (1996), a naming repertoire is demonstrated, for example, when a child acquires a novel tact following listener training alone. In the present

study, the participants typically did not emit foreign tacts at criterion level following standard listener training, but their typical performance of three to four correct responses was above chance level (with six comparison stimuli, the probability of three of six correct responses is 0.05). Although these data were not formally analyzed, it was noted that in tact and intraverbal probes following standard listener training, all participants made correct responses repeatedly with the same stimuli, and incorrect responses repeatedly with the same stimuli. This suggests that a subset of vocal relations were firmly in their repertoire as a result of listener training alone.

In this context, it may be noted that several studies conducted with children with developmental disabilities have found an effect of echoic response requirements during listener training on the acquisition of the listener repertoire, without measuring effects on derived vocal responding (Charlop 1983; Koegel et al. 1981; Leung and Wu 1997). A possible reason for this enhanced effect on acquisition is that the echoic response requirement may serve as a differential observing response (e.g., Dube and McIlvane 1999) to the dictated sample stimuli. In the present study, effects on acquisition were assessed in Experiment 2, and the data suggested no effect of the echoic response requirement on acquisition. If the participants were already echoing the dictated sample stimulus at the covert level, then again, a requirement to make this response overtly would not be expected to enhance acquisition.

If the participants already had a strong naming repertoire as defined by Horne and Lowe (1996), why did they not perform at criterion level when vocal relations were tested following listener training? In each experiment, six listener relations were taught concurrently. It seems possible that at the time the mastery criteria were met, some relations were more firmly in the participants' repertoire than others. Perhaps for some stimuli, the participants had a longer history of selecting the correct visual stimulus in the presence of the Japanese word, resulting in greater control by the stimuli involved in the task over collateral responses. Future research might address this issue by employing separate mastery criteria for each relation, and documenting at which point (before, during, or following attainment of that criterion) vocal tacts and intraverbals begins to emerge.

Although Experiment 1 addressed what we thought were the major procedural limitations of Experiment 1, it is still possible that the negative findings are related to some other procedural issue, rather than to participant characteristics. For example, Horne et al. (2004) reported that Bell (1999) observed the effect only after participants were required to echo the dictated names while looking directly at the relevant visual stimuli. In the present study, the participants were required to make the echoic response before the visual comparison stimuli were presented. As a result, they were not yet looking at the relevant visual stimulus when they made the echoic response. It is possible that a greater effect on tact emergence would

have been observed if the participants had been required to make the echoic response while touching the correct comparison. Alternatively, some other unidentified procedural variable may have prevented the demonstration of an effect in the present study.

In Experiment 1, all participants passed foreign-language tact probes following Phase 1 of exemplar training with only two of the six stimuli, and two participants also passed native-foreign intraverbal probes. Experimental control was demonstrated over the effect on tacts, as no participant met criterion for these two relations until exemplar training had been initiated, regardless of the amount of CRT and subsequent probe sessions that they had previously been exposed to. Because criterion performance with all six stimuli was achieved following direct training with only a subset of the stimuli, these results might seem consistent with an effect of multiple-exemplar training on the emergence of untrained vocal relations. However, the results should be interpreted with caution, as all participants were performing well above chance levels before exemplar training began. An analysis of probe performance with the stimuli that were not included in Phase 1 of exemplar training (data available from first author upon request) failed to suggest that Phase 1 generated increases in correct responses for those stimuli, with the exception of foreign tacts for Ashley. That is, the effect of exemplar training may be better accounted for as a direct training effect than an effect on relations that had not yet been trained.

A possible practical implication of the present data is that, although the inclusion of collateral echoic response requirements in listener training may facilitate acquisition (Charlop 1983; Koegel et al. 1981; Leung and Wu 1997) or facilitate the emergence of vocal responding (Ezell and Goldstein 1989; Hawkins et al. 2009) for children with language delays due to developmental disabilities, they may not do so for typically developing children who already have extensive verbal repertoires. Instead, other types of interventions might be considered for this population if listener training fails to produce highly accurate vocal tacts and intraverbals.

Acknowledgments We thank Dai Doan, Tereza Hopewell, Su Young Jung, Claire Kirk, Joel Kronawitter, Robert Pittman, Kesli Urbanczyk, and Sarah Willcox for assisting with data collection.

References

- Charlop, M. H. (1983). The effects of echolalia on acquisition and generalization of receptive labeling in autistic children. *Journal of Applied Behavior Analysis, 16*, 111–126.
- Connell, P. J., & McReynolds, L. (1981). An experimental analysis of children's generalization during lexical learning: Comprehension or production. *Applied Psycholinguistics, 2*, 309–332.
- Dube, W. V., & McIlvane, W. J. (1999). Reduction of stimulus overselectivity with nonverbal differential observing responses. *Journal of Applied Behavior Analysis, 32*, 25–33.
- Ezell, H. K., & Goldstein, H. (1989). Effects of imitation on language comprehension and transfer to production in children with mental retardation. *Journal of Speech and Hearing Disorders, 54*, 49–56.
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children severe language delays*. Boston: Allyn & Bacon.
- Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdes, C. (2005). 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.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum.
- Hawkins, E., Kingsdorf, S., Charnock, J., Szabo, M., & Gautreaux, G. (2009). Effects of multiple exemplar instruction on naming. *European Journal of Behavior Analysis, 10*, 265–273.
- Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 185–241.
- Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior, 81*, 267–288.
- Joyce, B. G., & Joyce, J. H. (1993). Using stimulus equivalence procedures to teach relationships between English and Spanish words. *Education and Treatment of Children, 16*, 48–65.
- Koegel, R. L., Dunlap, G., Richman, G. S., & Dyer, K. (1981). The use of specific orienting cues for teaching discrimination tasks. *Analysis and Intervention in Developmental Disabilities, 1*, 187–198.
- Lee, V. L. (1981). Prepositional phrases spoken and heard. *Journal of the Experimental Analysis of Behavior, 35*, 227–242.
- Leung, J., & Wu, K. (1997). Teaching receptive naming of Chinese characters to children with autism by incorporating echolalia. *Journal of Applied Behavior Analysis, 30*, 59–68.
- Luciano, C., Gomez Becerra, I., & Rodriguez Valverde, M. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior, 87*, 349–365.
- Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis, 28*, 115–126.
- Melchiori, L. E., de Souza, D. G., & de Rose, J. C. (2000). Reading, equivalence, and recombination of units: A replication with students with different learning histories. *Journal of Applied Behavior Analysis, 33*, 97–100.
- Petersdottir, A. I., Carr, J. E., Lechago, S. A., & Almason, S. M. (2008a). An evaluation of intraverbal training and listener training for teaching categorization skills. *Journal of Applied Behavior Analysis, 41*, 53–68.
- Petersdottir, A. I., & Hafliadottir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. *Journal of Applied Behavior Analysis, 42*, 685–690.
- Petersdottir, A. I., Olafsdottir, A. R., & Aradottir, B. (2008b). The effects of tact training and listener training on the emergence of bidirectional intraverbal relations. *Journal of Applied Behavior Analysis, 41*, 411–415.
- Rosales, R., Rehfeldt, R. A., & Lovett, S. (2011). Effects of multiple exemplar training on the emergence of derived relations in preschool children learning a second language. *The Analysis of Verbal Behavior, 27*, 61–74.

- Sidman, M., Wilson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities*, 6, 1–19.
- Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children*, 8, 67–76.
- Skinner, B. F. (1957). *Verbal behavior*. Acton: Copley.
- Sundberg, M. L. (2008). *Verbal behavior milestones assessment and placement program: The VB-MAPP*. Concord: AVB Press.
- Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions*, 18, 245–26.