Journal of Applied Behavior Analysis

JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2016, **49,** 1–6 NUMBER 3 (FALL)

EMERGENT VERBAL BEHAVIOR IN PRESCHOOL CHILDREN LEARNING A SECOND LANGUAGE

RICHARD J. MAY AND RACHEL DOWNS

UNIVERSITY OF SOUTH WALES

AMANDA MARCHANT

UNIVERSITY OF SOUTH WALES AND SWANSEA UNIVERSITY

AND

SIMON DYMOND

SWANSEA UNIVERSITY AND REYKJAVIK UNIVERSITY

We evaluated the emergence of untaught second-language skills following directly taught listener and intraverbal responses. Three preschool children were taught first-language (English) listener responses (e.g., "Point to the horse") and second-language (Welsh) intraverbal responses (e.g., "What is horse in Welsh?" [ceffyl]). After intervention, increases in untaught second-language tacts (e.g., "What is this in Welsh?" [ceffyl]) and listener responses (e.g., "Point to the ceffyl") were observed for all 3 participants.

Key words: derived relational responding, emergent relations, emergent tacts, second-language instruction, verbal behavior

Approximately two thirds of the world's population is either bilingual or multilingual (Baker, 2001). In the United Kingdom, promoting second language acquisition is an important goal for its linguistically diverse communities, and Welsh government policy requires teaching of both Welsh and English in primary and secondary schools.

Derived relational responding (DRR) is an effective and efficient approach for second-language instruction because only a subset of relations need be taught to evoke a range of untaught skills (Barnes & Rehfeldt, 2013). In a study with preschool children, Rosales, Rehfeldt, and Lovett (2011) successfully implemented second-language listener training (i.e., identifying objects given foreign language object names) to produce emergent second-language tacts (i.e., saying object names given objects). With adults, Dounavi (2014) found that teaching second-language intraverbals resulted in the emergence of second-language tacts and vice versa. In contrast, Petursdottir and Hafliðadóttir (2009) reported inconsistent results of DRR-based training with two 5-year-old children. Using a pretest–posttest design, they taught second-language intraverbals (i.e., saying the second-language name when hearing the native language name) and listener relations, but untaught tacts and intraverbals did not occur consistently at criterion levels.

Two limitations of the Petursdottir and Hafliðadóttir (2009) study might explain the discrepant findings. First, they did not assess maintenance of taught relations during tests for emergent relations. It is possible that the lack of emergent responding was due to failure to maintain the directly taught relations. Second,

We thank Caio Miguel and three anonymous reviewers for their helpful comments on an earlier version of this manuscript and Bryony Hawkins for her assistance with data collection.

Address correspondence to Richard J. May, University of South Wales, Pontypridd, CF37 IDL, United Kingdom (e-mail: Richard.may@southwales.ac.uk).

doi: 10.1002/jaba.301

the absence of programmed reinforcement during test phases may have disrupted on-task behavior (cf. LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003). The present study sought to address these concerns by including reinforced maintenance probes in test sessions. In addition, we implemented a multiple baseline design across behaviors, which confers greater experimental control than pretest–posttest designs.

METHOD

Participants and Setting

Three typically developing children participated. Alice was a 52-month-old girl, Stew was a 48-month-old boy, and Betty was a 30-month-old girl. Participants' English language skills were assessed with the British Picture Vocabulary Scale (2nd ed., Dunn, Dunn, Whetton, & Burley, 1997) and the Expressive One-Word Picture Vocabulary Test (Martin & Brownell, 2011). All participants scored within the normal range on both assessments. Welsh language skills were not formally assessed; however, none of the participants spoke Welsh at home, and teachers reported that their vocabularies were limited to a small number of common nouns and adjectives (i.e., common items, colors). Before the study, participants were able to echo all of the words included in the stimulus sets. Participants were not provided with additional Welsh language instruction involving the experimental stimuli during the study.

Sessions with Alice and Stew were conducted in a preschool, and sessions with Betty were conducted in her home. During sessions, the child sat at a table and the experimenter sat at a 90° angle to the child.

Stimuli, Dependent Variables, and Interobserver Agreement

We used two six-item stimulus sets with each participant. Stimuli included English spoken words (Set 1: "horse" and "cow"; Set 2: "strawberry" and "cake"), Welsh spoken words (Set 1: "ceffyl" and "buwch"; Set 2: "mefys" and "teisen"), and their corresponding pictures. During pretraining, we used stimuli that consisted of items (e.g., animals) that the children could already name and identify in Welsh. Picture stimuli consisted of color illustrations that measured 8 cm by 12 cm. A sticker chart (30 cm by 21 cm) was also used.

The primary dependent variable was the percentage of tact-Welsh and listener-Welsh responses during pretest and posttest sessions. Tact-Welsh responses were assessed by asking participants to state the Welsh name when shown a picture. Listener-Welsh responses were assessed by asking participants to select a picture from an array of two comparisons when given the Welsh name. A correct response was scored when the child uttered the correct vocal response (or approximation) during tact trials or pointed to the correct picture during listener trials within 5 s of the instruction.

A second observer scored 60% of sessions (84% of test sessions, 32% of training sessions), either in person or from videotapes, to obtain interobserver agreement; agreement or disagreement between the two observers was determined for each response. Interobserver agreement was calculated for each session by dividing the number of agreements by the total number of trials and converting the result to a percentage. Agreement averaged 99.6% for Alice (range, 94% to 100%), 97% for Stew (range, 88% to 100%) and 99% (range, 94% to 100%) for Betty.

Design

A multiple baseline design across stimulus sets was used to evaluate effects of direct training on emergent responding. After criterion-level (≥75% correct) responding in posttest sessions for emergent relations with Set 1, tests for emergent responding with an additional set of stimuli (Set 2) were implemented.

Procedure

General procedure. During all conditions, participants were shown the sticker chart and asked if they would like to play a game to win a prize. During training, correct responses were followed by praise, and stickers were provided on a variable-ratio (VR) 3 schedule. Incorrect responses or nonresponses were followed with corrective feedback, including a prompt to engage in the correct response (e.g., "That is not the right answer; the right answer is [correct answer]") or pointing towards the correct comparison. Reinforcement was not provided after a correction. During pretest and posttest sessions, listener-Welsh and tact-Welsh trials never included reinforcement or correction. During listener trials, comparison stimuli consisted of two pictures from the same set. The order of trials was randomized, and the position of comparisons across trials was counterbalanced. Four to eight sessions were conducted once per day, approximately two to three times per week; each lasted approximately 5 min.

Pretraining. Here, tact-Welsh and listener-Welsh trials were presented in an identical fashion to the pretest phase, but with familiar stimuli. All responses were followed by feedback (either reinforcement or correction). Sessions consisted of eight trials, with two listener relations and two tact relations presented twice each. The criterion was 100% correct within a session. All participants passed the pretraining phase within two sessions.

Pretest. During this phase, tact-Welsh and listener-Welsh trials were presented in the absence of feedback. Each of the four relations was presented twice. In addition, eight motor-imitation trials were interspersed in test sessions. During these trials the experimenter asked, "Can you do this?" and presented a simple motor model (e.g., clapping). The experimenter implemented physical guidance if incorrect motor-imitation responses occurred and reinforced both correct and prompted responses. Participants had to score no more than five of eight (≤63%) correct on the combined tact- and listener-Welsh relations for that stimulus set to be used during Training 1.

Training 1 (listener training). The experimenter presented two pictures from the same stimulus set and said, "Point to the [English word]." Sessions consisted of eight trials, with each relation presented four times. The criterion to progress to Training 2 was seven of eight trials correct (87.5%) during one session.

Training 2 (intraverbal training). During Training 2, the experimenter asked, "What is [English word] in Welsh?" The number of trials and mastery criterion were identical to listener training.

Training 3 (mixed listener and intraverbal training). Training 3 interspersed Training 1 and Training 2 trial types. Sessions consisted of 16 trials, with each trial type presented four times. The criterion was set at 14 of 16 correct responses (87.5%).

Training 4 (mixed listener and intraverbal training: 50% feedback). Training 4 was identical to Training 3, with the exception that only 50% of trials were followed by feedback, to approximate the rate of reinforcement to be presented during the posttest phase. Reinforced and nonreinforced trials were randomly interspersed.

Posttest. This was identical to the pretest, except that taught relations were randomly interspersed with test trials. Correct responses during taught-relation trials were reinforced. The posttest was conducted in 16-trial sessions that involved two presentations of each trial type. Criterion was set at seven of eight trials correct for the taught relations. If a participant failed to achieve this criterion, he or she was returned to Training 3. A final posttest session was conducted to assess maintenance after 2 to 4 weeks, during which no training or testing occurred.

RESULTS AND DISCUSSION

Figure 1 illustrates pretest and posttest sessions. During the pretest sessions, participants responded to tact-Welsh relations at a mean accuracy of 0% to 25% correct for Set 1 and 0% to 19% for Set 2, and they responded to listener-Welsh relations at a mean accuracy of 37.5% to 100% for Set 1 and 42% to 50% for Set 2. Alice, Stew, and Betty completed Training 1 through Training 4 in 21, 10, and 13 sessions, respectively. Alice repeated Training 3 with Set 1 before the posttest in Session 5 because she had been unable to attend any sessions for a week. Otherwise, no remedial training was needed. As Figure 1 shows, the emergence of derived relations corresponded with the staggered implementation of training (with the exception of listener relations for Set 1 for Stew). After training with Set 1, Alice initially responded at 50% accuracy for both emergent relations (Session 3), which then rose to 75% from Session 5 through Session 9. Betty responded at or above 75% accuracy for Set 1 posttest probes, with the exception of emergent listener responses in Session 4 (50% correct). In all other posttest sessions, participants

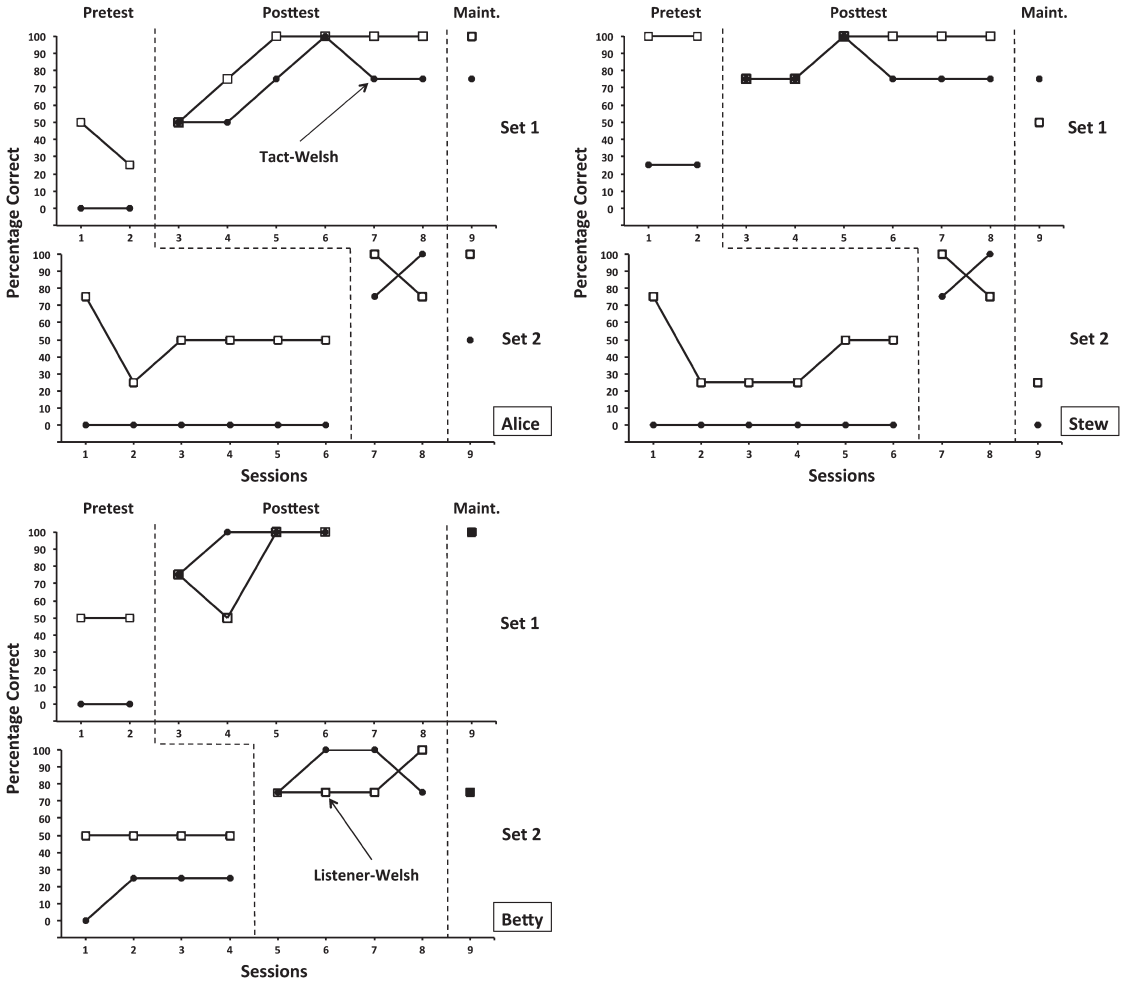

Figure 1. Percentage correct during pretest, posttest, and maintenance phases.

responded at or above 75% accuracy for both emergent repertoires. Maintenance posttests were conducted after 2-, 3-, and 4-week intervals for Alice, Stew, and Betty, respectively. With Alice and Betty, taught and emergent relations were largely intact. Stew's responding was below criterion accuracy for both the taught and emergent relations.

These findings indicate that teaching first-language listener and second-language intraverbal responses resulted in increases in untaught second-language tact and listener responses for all three participants. Follow-up test probes demonstrated that for two of the participants, emergent responses persisted for 2 to 4 weeks after the initial posttest. This latter finding is notable because measures of durability of emergent responding are often omitted from DRR studies (Rehfeldt, 2011).

The present study has some potential limitations. Although increases in emergent responding were observed, the effect was not entirely robust. Participants frequently responded at less than perfect accuracy during posttest sessions, despite near-perfect accuracy for the trained relations. The extent to which listener relations emerged for Stew after training with Set 1 can also be questioned. Although combined emergent relations were below the pretest criterion, the listener-Welsh relations were at 100% accuracy in pretests. A further limitation concerns the use of a two-stimulus array during listener trials (see Sidman, 1980). Future research should employ stimulus arrays of at least three stimuli. Another limitation is the lack of parity between the pretest and posttest conditions; motor-imitation targets were interspersed in pretests, whereas previously trained responses were used in posttests. The inclusion of taught relations during posttests might have provided additional cues for correct responding that were not present during pretests. Future work could mitigate these concerns by incorporating unrelated conditional discriminations into pretests as well as structuring the presentation of trained and test relations to avoid inadvertent cueing. Finally, incorporation of measures to ensure fidelity of the training and testing procedures is advised.

Overall, these findings suggest that DRR-based procedures are an effective way to establish second-language tacting and listener skills in typically developing children. These results differ from those of Petursdottir and Hafliðadóttir (2009), who found that teaching new intraverbal responses combined with existing listener behavior did not result in criterion-level emergence of derived tacting. Interspersal of reinforced maintenance trials may have accounted for the more favorable outcomes in the current study. Alternatively, the different outcomes may have occurred due to participant characteristics (e.g., language skills) or familiarity with the tasks involved. Our participants were from bilingual communities, and thus, may have been familiar with the contextual cues present in the questions. Future research should seek to elucidate further the factors that contribute to the success of DRR applications in young typically developing populations.

REFERENCES

Baker, C. (2001). *Foundations of bilingual education and bilingualism* (3rd ed.). Clevedon, UK: Multilingual Matters.

Barnes, C. S., & Rehfeldt, R. A. (2013). Advances in language interventions based on relational frame theory for individuals with developmental disorders. In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory: Research and application* (pp. 151–176). Oakland, CA: New Harbinger.

Dounavi, K. (2014). Tact training versus bidirectional intraverbal training in teaching a foreign language. *Journal of Applied Behavior Analysis*, *47*, 165–170. doi: 10.1002/jaba.86

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd ed.). London, UK: Nelson.

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children

diagnosed with autism. *Behavioral Interventions*, *18*, 279–289. doi: 10.1002/bin.144

Martin, N., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test* (4th ed.). Novato, CA: ATP Assessments.

Petursdottir, A. I., & Haflidadóttir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. *Journal of Applied Behavior Analysis*, *42*, 685–690. doi: 10.1901/jaba.2009.42-685

Rehfeldt, R. A. (2011). Toward a technology of derived stimulus relations: An analysis of articles published in the *Journal of Applied Behavior Analysis*, 1992–2009. *Journal of Applied Behavior Analysis*, *44*, 109–119. doi: 10.1901/jaba.2011.44-109

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. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior, 33*, 285–289. doi: 10.1901/jeab.1980.33-285

Received February 26, 2015
Final acceptance November 20, 2015
Action Editor, Einar Ingvarsson