

Emergent Verbal Behavior in Preschool Children Learning a Second Language

Richard J. May^a, Rachel Downs^a, Amanda Marchant^{a,b}, and Simon Dymond^{c,d}

^aSchool of Psychology, University of South Wales, Pontypridd, CF37 1DL, United Kingdom.

^bCollege of Medicine, Swansea University, Swansea, SA2 8PP, United Kingdom.

^cDepartment of Psychology, Swansea University, Swansea, SA2 8PP, United Kingdom.

^dDepartment of Psychology, Reykjavík University, Menntavegur 1, Nauthólsvík, 101
Reykjavík, Iceland.

doi: 10.1002/jaba.301

This manuscript is a Pre-publication version.

This article may not exactly replicate the final version published in the SEAB journal. It is not
the copy of record. For the final published version please see

[http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1938-3703](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1938-3703)

Address correspondence to: Richard J. May, University of South Wales, Pontypridd, CF37
IDL, United Kingdom. Email: Richard.may@southwales.ac.uk

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”]). Following the intervention, increases in untaught second-language facts (e.g., “What is this in Welsh?” [“Ceffyl”]) and listener responses (e.g., “Point to the ceffyl”) were observed for all three participants.

Emergent Verbal Behavior in Preschool Children Learning a Second Language

Approximately two thirds of the world's population are either bilingual or multilingual (Baker, 2001). Within 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 school.

Derived relational responding (DRR) is an effective and efficient approach for second language instruction, because only a subset of responses 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., stating object names given objects). With adults, Douvani (2014) found that teaching second-language intraverbals resulted in the emergence of second-language tacts, and vice versa.. In contrast, Pétursdóttir and Hafliðadóttir (2009) reported inconsistent effects of DRR-based training in facilitating untaught verbal relations with two five-year old children. Using a pretest/posttest design, they taught second-language intraverbals (i.e., stating 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 Pétursdóttir and Hafliðadóttir (2009) study might explain the discrepant findings. First, the authors 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, the absence of programmed reinforcement during extensive 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. Additionally, we

implemented a multiple baseline across behaviors design, 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 48-month-old-boy, and Betty was a 30-month-old girl. Participants' English language skills were assessed with the *British Picture Vocabulary Scale-Second Edition* (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 teacher reports indicated their vocabularies were limited to a small number of common nouns and adjectives (i.e., common items, colors). Prior to the study, participants were able to echo all of the words included in the stimulus sets.

Sessions with Alice and Stew were conducted in a preschool setting, while 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. Participants were not provided with additional Welsh language instruction involving the experimental stimuli during the study.

Experimental stimuli, dependent variables, and inter-observer 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 consisting of items (e.g., animals) that the children could already name and identify in Welsh. Picture stimuli consisted of color illustrations measuring 8 cm x 12 cm. A sticker chart measuring 30 cm x 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 an animal (or food) 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 comparison stimulus during listener trials, within 5 s of the instruction.

A second observer scored 60% percent of sessions (84% of test sessions, 32% of training sessions) either in person or from videotapes to obtain interobserver agreement (IOA); agreement or disagreement between the two observers was determined for each response. IOA was calculated for each session by dividing the number of agreements by the total number of trials and multiplying by 100. IOA averaged 99.6% for Alice (range 94 - 100%), 97.5% for Stew (range 88 -100%) and 98.8% (range 94 - 100%) for Betty.

Experimental Design

A multiple baseline across stimulus sets was used to evaluate effects of direct training on emergent responding. Following criterion level ($\geq 75\%$ 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 procedures. 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 3 (VR3) schedule. Incorrect responses or non-responses 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 following a correction. During *Pretest* and *Posttest* sessions, Listener-Welsh and Tact-Welsh trials were never followed by reinforcement or correction. During listener trials, comparison stimuli consisted of two picture stimuli from the same set. The order of trial presentations was randomized, and the position of comparisons across trials was counterbalanced. Between four to eight sessions were conducted once per day, approximately two to three times per week, with each session lasting around 5 minutes.

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 were 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 out of eight ($\leq 63\%$;) correct on the combined Tact- and Listener-Welsh relations for that stimulus set to be used during *Training 1*.

Training 1 (Listener training). Here, the experimenter presented two pictures from the same stimulus set and the instruction, “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 out 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 was identical to listener training.

Training 3 (Mixed listener and intraverbal training). *Training 3* interspersed *Training 1* and *Training 2* trial types. Sessions consisted of sixteen trials, with each trial type presented four times. The criterion was set at fourteen out of sixteen 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, in order to approximate the rate of reinforcement to be presented during the *Posttest* phase. Reinforced and non-reinforced 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 sixteen-trial sessions, involving two presentations of each trial type. Criterion was set at seven out of eight trials correct for the taught relations. If the participants failed to achieve this criterion during a *Posttest* session, they were returned to *Training 3*. A final *Posttest* session was conducted to assess maintenance following a period of between two and four weeks, during which no training or testing was conducted.

Results and Discussion

Figure 1 illustrates *Pretest* and *Posttest* sessions. Participants responded to Tact – Welsh relations at a mean accuracy of between 0% and 25% correct for Set 1 and between 0% and 19% for Set 2, and to Listener – Welsh relations at a mean accuracy of between 37.5% and 100% for Set 1 and between 42% and 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 prior to *Posttest* 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 occurred in accordance with the staggered implementation of training (with the exception of listener relations for Set 1 for Stew). Following 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 responded at or above 75% accuracy for both emergent repertoires. Maintenance *Posttests* were conducted following a two-, three-, or four week interval 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 between two and four weeks following the initial posttest period. This latter finding is notable because measures of durability of emergent responding are often omitted from studies on DRR (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 (as low as 50%) during *Posttest* sessions despite near perfect accuracy for the trained relations. The extent to which listener relations emerged for Stew following 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 pretest. 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, incorporating measures to ensure fidelity of the training and testing procedures is advised.

Overall, these findings suggest that employing DRR-based procedures is an effective way of establishing second-language tacting and listener skills in typically developing children. These results differ from those of Pétursdóttir 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 differing 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 further elucidate the factors that contribute to the success of DRR applications in young typically developing populations.

Author note

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.

References

- Barnes, C. S., & Rehfeldt, R. A. (2013). Advances in language interventions based on Relational Frame Theory for individuals with developmental disorders. In S. Dymond and B. Roche (Eds.), *Advances in relational frame theory: Research and application* (pp. 151 -176). Oakland, CA: New Harbinger.
- Baker, C. (2001). *Foundations of bilingual education and bilingualism* (3rd ed.). Clevedon, UK: Multilingual Matters.
- 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 (II)*. London: 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 – Fourth Edition (EOWPVT-IV)*. Novato, CA: ATP Assessments.
- Pétursdóttir, A. I., & Hafliðadóttir, R. D. (2009). A comparison of four strategies for teaching a small foreign language vocabulary. *Journal of Applied Behavior Analysis*, 42, 497-745. doi: 10.1901/jaba.2009.42-685
- Rehfeldt, R. A. (2011). Toward a technology of derived stimulus relations: An analysis of articles published in JABA, 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 discriminations. *Journal of the Experimental Analysis of Behavior*, 33, 285–290. doi: 10.1901/jeab.1980.33-285

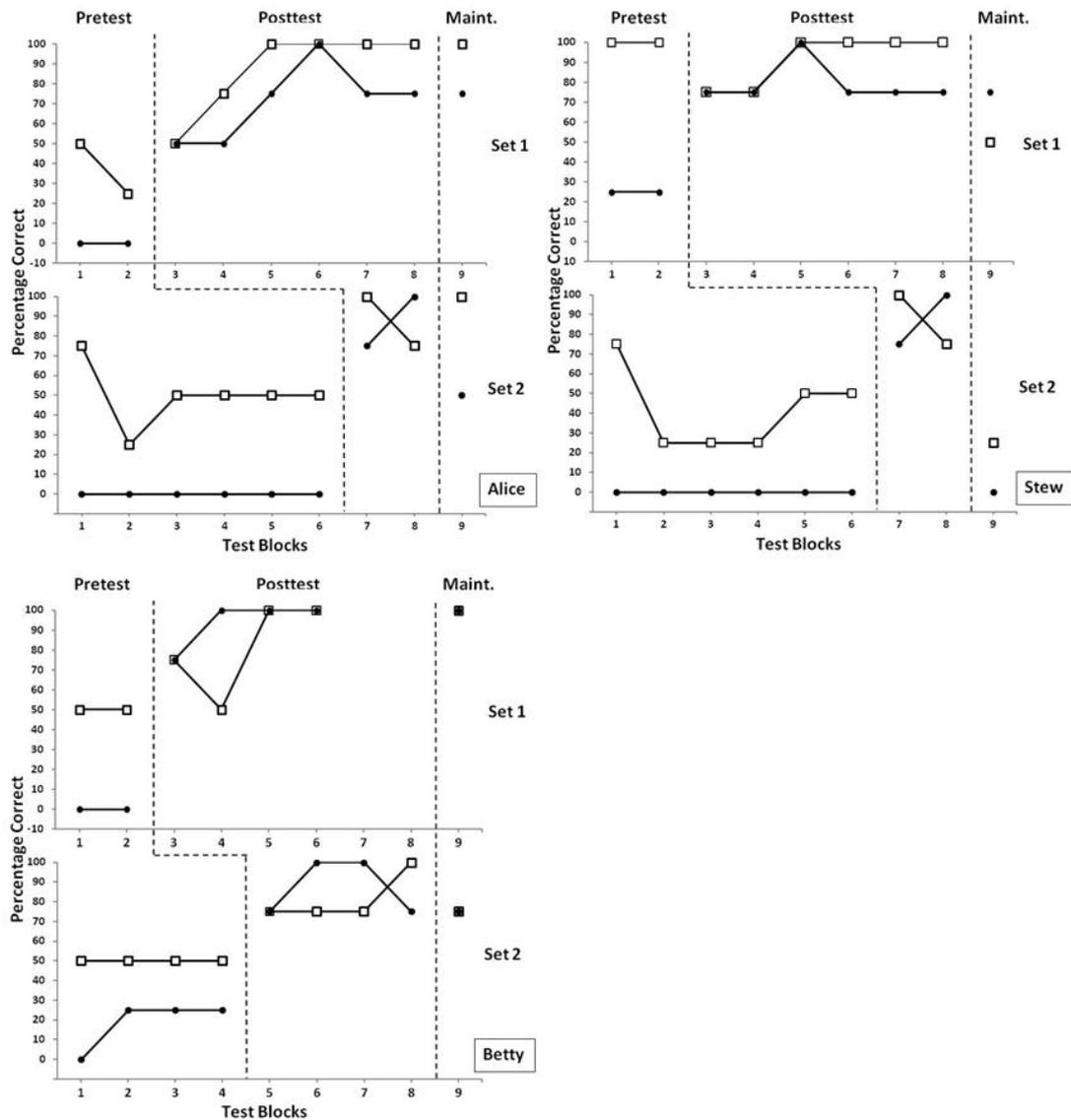


Figure 1. Percentage correct during Pretest, Posttest, and Maintenance (“Maint.”) phases.

Circle markers indicate Tact-Welsh relations. Open square markers indicate Listener-Welsh relations.