Western Michigan University
ScholarWorks at WMU

Dissertations Graduate College

8-2006

An Evaluation of Intraverbal Training and Listener Training for Teaching Categorization Skills

Anna Ingeborg Petursdottir
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Clinical Psychology Commons, Educational Psychology Commons, and the School Psychology Commons

Recommended Citation

Petursdottir, Anna Ingeborg, "An Evaluation of Intraverbal Training and Listener Training for Teaching Categorization Skills" (2006). *Dissertations*. 977.
https://scholarworks.wmich.edu/dissertations/977

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.

AN EVALUATION OF INTRAVERBAL TRAINING AND LISTENER TRAINING
FOR TEACHING CATEGORIZATION SKILLS

by

Anna Ingeborg Petursdottir

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fullfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
August 2006

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

AN EVALUATION OF INTRAVERBAL TRAINING AND LISTENER TRAINING FOR TEACHING CATEGORIZATION SKILLS

Anna Ingeborg Petursdottir, Ph.D.

Western Michigan University, 2006

Behavioral language interventions, such as those employed in early and intensive behavioral intervention (EIBI) programs, target both expressive and receptive language skills. Skinner's (1957) analysis of verbal behavior provides a framework for analyzing expressive and receptive language in terms of stimulus control and reinforcement history. From this perspective, different expressive language programs target different verbal operants, such as *tacts*, *intraverbals*, and *echoics*, whereas most receptive language programs target a type of listener behavior that may be referred to as *manded stimulus selection* (Michael, 1995). Although EIBI curricula (e.g., Maurice, Green, & Luce, 1996; Leaf & McEachin, 1999) have frequently recommended teaching receptive before expressive skills, the empirical literature suggests that the reverse sequence may sometimes be more efficient. Specifically, tact training may be more likely to generate an emergent listener repertoire than listener training to generate an emergent tact repertoire. Less is known about the extent to which a similar relation holds for intraverbals and listener behavior, even though the sequencing of intraverbal and listener training is a consideration in many language training programs, such as those that teach various categorization skills.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The purpose of the present study was to provide an evaluation of both intraverbal and listener categorization training on untrained categorization skills. The evaluation was conducted with six typically developing 3-year-old children, who learned to categorize previously unfamiliar stimuli (i.e., characters from foreign writing systems, outline maps of foreign countries). No emergence of untrained categorization skills was observed among three children who received listener training, and untrained skills did not emerge reliably among three other children who received intraverbal training. These results are in line with the notion, suggested by Skinner's (1957) analysis, that listener and speaker repertoires may be functionally independent of one another. From an applied perspective, additional research is needed on how to most efficiently sequence language training programs that incorporate the training of both intraverbal and listener relations, and basic research on the establishment of topography-based verbal relations also appears warranted.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3234887

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3234887

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

© 2006 Anna Ingeborg Petursdottir

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGEMENTS

I would like to thank my advisor and dissertation committee chair, Dr. James E. Carr, for his outstanding mentorship, and for his input and recommendations regarding the present project. I also thank committee members Dr. Jack Michael, Dr. Scott Gaynor, Dr. Cynthia Pietras, and Dr. Norm Peterson for taking the time to review my work and providing helpful comments. Additionally, I thank Dr. Michael for being available to discuss and clarify aspects of Skinner's analysis of verbal behavior. I would also like to acknowledge the contributions of my friend and colleague, Dr. Caio Miguel, who gave me the opportunity to collaborate on previous projects that partially inspired the present work. Our collaboration and long conversations on verbal behavior and derived stimulus relations were crucial in providing me with the conceptual background necessary to pursue this project.

I am greatly indebted to Sarah A. Lechago for overseeing data collection while I was away at the University of Akureyri in Iceland from August 2005 to May 2006. Without her gracious contribution of time and effort, this project would not have been completed in a timely manner. I thank Season M. Almason for providing considerable assistance with data collection, as well as a number of undergraduate students who served as research assistants on this project: Kelsey Behnke, Heather DeForest, Allison DeGraaf, Juliet Flynn, Adam Freeman, Jeana Koerber, Tracy Lepper, Catrina Litzenburg, Sadie Lovett, Jada Rae-Traub, Stacy Ruse, and Kelly Stone. I also thank

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Sophie Rubin for assisting with data collection on several occasions during winter break when regular staff were unavailable.

I thank the director and staff at the child-care center where the study was conducted for their patience and support, and for the provision of a pleasant work environment. My warmest acknowledgements go to the children who participated in the study and made all the effort worthwhile, and I thank the children's parents for giving permission for them to participate.

A number of others deserve acknowledgement: The University of Akureyri for paying for some of my time on this project; Dr. Z. Gabriela Sigurdardottir for sparking my interest in behavior analysis and sending me off to graduate school; numerous other members of ICEABA for inspiring interactions through the years; all of my friends for their company, support and interest in my work; in particular, Barbara and John Esch; and the Fourth Coast Coffee Shop and Coffee Works, of Kalamazoo, MI, for providing stimulating writing environments and serving much-needed caffeinated beverages.

Lastly, I thank my late grandfather, Ólafur Sigurðsson, for providing me with respect for higher education and for demonstrating every day how it can be put to use for the benefit of others, and I thank my entire family in Iceland for their support, as well as for putting up with my absence and my preoccupation with graduate school in recent years.

Anna Ingeborg Petursdottir

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

INTRODUCTION

Sequencing of Programs in Behavioral Language Interventions

For decades, behavior-analytic techniques have been utilized for teaching verbal skills to individuals with language impairments. The most comprehensive texts that have been published on behavioral language interventions are ones that describe programs or procedures for teaching language to young children with autism, within the context of early and intensive behavioral intervention (*EIBI*) programs (Leaf & McEachin, 1999; Lovaas, 2003; Maurice, Green, & Luce, 1996; M. L. Sundberg & Partington, 1998). EIBI texts vary from one another with respect to the specific teaching strategies that they recommend. However, all are similar in that they recommend breaking down complex language skills into small steps and teaching them in a structured format, using such techniques as positive reinforcement, discrimination training, and transfer-of-stimulus-control procedures. Further, the overall content of beginning and intermediate language curricula is similar across texts, and generally resembles the content of a curriculum originally developed by O. Ivar Lovaas and his colleagues at UCLA. This curriculum and its development was described in detail in early publications (Lovaas, 1977, 1981) and used in a landmark outcome study (Lovaas, 1987) on the effects of EIBI on intellectual functioning and school placement. In the curriculum, children are initially taught to respond to simple instructions, imitate the speech of others, and identify common stimuli in their environment. Once these basic skills are mastered, more complex skills are introduced. For example, the children are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

taught to speak in complete and grammatical sentences, describe details of their environment, participate in conversation, and tell stories.

A potentially under-explored area of research on behavioral language interventions concerns the optimal sequencing of programs that target different types of skills, such that minimal teaching effort will result in maximum gains. In the original UCLA curriculum, the recommended sequence of programs was based partially on the perceived complexity of the discriminations involved, and partially on the order in which typically developing children had been observed to acquire various language skills (Lovaas, 1977). A similar sequence of programs is still presented in contemporary EIBI curricula (Leaf & McEachin, 1999; Maurice et al., 1996). However, Lovaas (1977, 2003) has cautioned that a typical developmental sequence may not always be the ideal teaching sequence, and that the sequencing of programs ultimately needs to be based on empirical data. Of course, certain skills and skill components must logically precede others; for example, a child must be able to emit speech sounds before he or she can be taught to vocally name objects. Other types of skills, however, impose no such logical constraints on sequencing, and thus the question of optimal sequencing becomes an empirical one.

One question regarding sequencing that has long concerned professionals implementing language interventions (Watters, Wheelers, & Watters, 1981) is whether *receptive* or *expressive* language skills should be taught first. Lovaas (1977) tentatively recommended completing receptive programs before introducing the corresponding expressive programs; for example, teaching children to point to objects upon hearing their names before teaching them to name those same objects. This recommendation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was reportedly based in part on clinical data indicating that acquisition of the former could sometimes facilitate the acquisition of the latter, and was also consistent with developmental observations indicating that in typical language development, language comprehension tends to precede production (e.g., Fraser, Bellugi, & Brown, 1963). Consequently, at least two of the major contemporary EIBI manuals (Leaf & McEachin, 1999; Maurice et al., 1996) typically list the mastery of a receptive program as a prerequisite for the introduction of any related expressive programs. Neither manual, however, cites empirical data in support of this sequence. More recently, Lovaas (2003) has suggested that this recommendation perhaps should no longer be upheld, following the publication of an experimental study (Wynn & Smith, 2003) which indicated that some children diagnosed with autism might benefit from being taught expressive skills first.

The terms expressive and receptive language, however, both encompass a variety of skills that differ from one another in terms of stimulus control and behavioral function. A behavioral analysis of these skills, along with experimentation designed to uncover their relations with one another during acquisition, may result in findings that shed light on optimal teaching sequences.

A Behavioral Conceptualization of Expressive and Receptive Language

From a behavioral perspective, expressive and receptive language skills do not represent different manifestations of the same underlying linguistic knowledge. Rather, the two types of skills can be understood as instances of different types of stimulus control over topographically distinct responses. Skinner's (1957) analysis of verbal behavior provides a useful framework for analyzing expressive and receptive language

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in terms of stimulus control and reinforcement history. Some EIBI providers, in fact, prefer the use of this framework for describing their programs to the more widely used expressive/receptive language distinction (e.g., M. L. Sundberg & Partington, 1998). Skinner distinguished between the behavior of the speaker, which may be considered roughly analogous to expressive language, and the behavior of the listener, which is similar to receptive language. The behavior of the speaker, according to Skinner, is *verbal* behavior. Any behavior may be considered verbal if it fulfills the requirement of having a history of reinforcement mediated by another person, who has been "conditioned precisely in order to reinforce the behavior of the speaker" (p. 225). Behaving as a listener, on the other hand, involves responding to a verbal stimulus in a way that is "not necessarily verbal in any special sense" (p. 2). In other words, the verbal stimulus evokes a response that has been reinforced by its effects on the nonsocial environment without the mediation of another person's behavior.

Skinner's (1957) primary emphasis was on the behavior of the speaker. In his analysis, he subdivided the speaker's behavior into a number of verbal operants, on the basis of reinforcement history and the resulting type of stimulus control over the form or topography of the response. Among the basic verbal operants that Skinner described are the *mand*, the *tact*, the *echoic*, and the *intraverbal*. In the mand relation, the response topography is controlled by a motivational variable, as a result of a history of being reinforced with a specific consequence that is effective as a reinforcer in the presence of that variable. For example, if the response form "water" is reinforced with water under conditions of water deprivation, water deprivation subsequently tends to evoke that response as a mand. By contrast, in the tact, echoic and intraverbal relations,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the form of the response is under discriminative rather than motivational stimulus control, as a result of a differential reinforcement history involving many different reinforcers or generalized conditioned reinforcers. In the case of the tact, the discriminative stimulus (S^D) is nonverbal. Examples of tacts include the vocal response "cat" evoked by the sight of a cat, and the same response evoked by the sound of a cat purring. In EIBI curricula, most expressive language programs that teach "naming" can be characterized as tact training. In the echoic relation, the S^D is a verbal stimulus; that is, a stimulus produced by another person's verbal behavior. Further, there is point-to-point correspondence between parts of the verbal stimulus and parts of the response that it controls, such that according to Skinner "the response generates a sound-pattern similar to that of the stimulus" (p. 55). Repeating any word, such as "cat", as a result of hearing someone else say it, is an example of an echoic relation, and in EIBI curricula, the establishment of an accurate echoic repertoire is the objective of programs that teach verbal imitation. Finally, in the intraverbal relation, the S^D is also verbal, but does not have point-to-point correspondence with the response. Examples of intraverbals may include appropriate responding to such verbal stimuli as /What is your name?/, /What does a cat say?/, or /Name some animals./. Intraverbal training is included in a variety of intermediate and advanced EIBI programs, including ones that target social interactions and conversational skills, as well as general knowledge.

Most EIBI programs that teach receptive skills may be usefully considered to teach listener rather than speaker behavior, because although reinforcement may be mediated by the teacher during training, the ultimate goal is to train responses to verbal stimuli that will eventually be maintained by their effects on the nonsocial environment.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As with the behavior of the speaker, it is possible to subdivide the behavior of the listener into different types of operants (Michael, 1996). It may be useful to distinguish between two types of listener relations commonly addressed in EIBI programs, based on the type of discrimination required of the listener. One type of listener relation requires only a simple discrimination. The listener makes a response, the topography of which is specified by the speaker; for example, a child sits down as a result of being asked to /Sit down./ This type of behavior is the target of EIBI programs that teach the child to follow simple instructions. The second type of listener relation involves a conditional discrimination, in which a stimulus-selection response, such as pointing to or picking up an object, is evoked by the sight of the object (the S^D) in the presence of a specific verbal stimulus (the *conditional stimulus*); for example, the spoken name of the object. This type of behavior has sometimes been referred to as manded stimulus-selection (Michael, 1985), whereas the first type might be referred to as simple mand compliance. As an example of manded stimulus-selection, the verbal stimulus /Where is the cat?/ might establish the sight of a cat as an S^D for a pointing response. In EIBI curricula, this type of listener behavior is the target of all programs that teach receptive identification of objects and attributes, by training the child to touch or otherwise select one stimulus from an array of stimuli in response to a verbal stimulus. A single program may require both types of listener behavior simultaneously, when the verbal stimulus delivered by the teacher specifies both the stimulus to be selected and the topography of the selection response. For example, in a program that teaches prepositions, a child may be taught to respond to an instruction such as /Put the block in front of the box./, which specifies the selection of a particular location (in front of the box), as well as the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

topography of the selection response (block placement). However, only one of these relations is typically the primary target of each program; in the case of prepositions, for example, the selection of the specified location is the critical target repertoire.

It is important to note that according to Skinner (1957), there is no requirement for the response topography in a verbal operant to be vocal. The topography may also consist, for example, of a sign, a gesture, typing or writing, or pointing to a verbal stimulus, such as a written word. The function of the operant is the same regardless of its topography. It has been pointed out, however, that it is possible to distinguish between two types of verbal operants on the basis of whether they are *topography-based* or *selection-based* (Michael, 1985). In a topography-based verbal relation, according to Michael, a particular stimulus class evokes a particular response topography that is distinct from topographies evoked by other stimulus classes. For example, the sight of a cat may evoke the response "cat" as a tact, whereas the sight of a dog evokes the topographically different response "dog". A selection-based verbal relation, by contrast, is similar to listener relations of the manded stimulus-selection type, in that a conditional stimulus enhances the control of an S^D over a selection response, such as pointing. The critical difference is that in a listener relation the conditional stimulus is verbal and the S^D nonverbal, whereas in a verbal relation, the selected S^D is necessarily verbal while the conditional stimulus can be either verbal or nonverbal. In selection-based relations, the topography of the response does not covary with the controlling stimulus class. The sight of a cat may evoke the response of pointing to the written stimulus CAT as a selection-based tact, and the sight of a dog may evoke a near-identical topography when the listener points to DOG. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

topography may differ depending on the location of the written stimulus, but does not vary systematically with the nonverbal stimuli that evoke it.

The distinction between selection-based and topography-based verbal behavior is considered by many to be of conceptual importance and to have implications for language-training programs (Hall & Chase, 1991; Michael, 1985; Polson & Parsons, 2000; Potter & Brown, 1997; M. L. Sundberg & Partington, 1998). Some of these considerations will be described in later sections. For now, it will be noted that EIBI programs typically focus primarily on teaching vocal language, which is topography-based. If children fail to acquire vocal language, sign-language or another topography-based behavior may be considered, or alternatively, a selection-based communication system (M. L. Sundberg & Partington, 1998). The major EIBI curricula, however, tend to describe their programs primarily with vocal language training in mind (Leaf & McEachin, 1999; Maurice et al., 1996).

Functional Independence of Speaker and Listener Behavior

According to Skinner's (1957) analysis, the behavior of the speaker and the behavior of the listener are functionally independent of one another. The two types of behavior consist of topographically different responses under the control of functionally dissimilar stimuli, requiring separate histories of reinforcement to be established. Therefore, acquiring appropriate listener behavior with respect to a verbal stimulus such as /Show me the cat./, should not necessarily result in, for example, the ability to say "cat" in the presence of cats. Neither should the acquisition of the vocal response "cat" as a tact or any other verbal operant result in the ability to select a cat upon hearing the spoken word. For the same reason, each of the verbal operants is functionally

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent from all others, such that, for example, reinforcement of the response "cat" as a tact should not simultaneously establish intraverbal control over the response "cat" by the question /Name an animal./. This notion of functional independence implies that beginning language learners, such as most young children with autism, may require explicit training of both listener behavior and the various verbal operants that comprise speaker behavior. From this perspective, the sequence in which receptive and expressive programs are introduced may not be of fundamental importance. Practically, of course, it would be immensely difficult to train vocal tact or intraverbal relations without first establishing an echoic repertoire that would enable the teacher to prompt vocal responses in the presence of relevant verbal or nonverbal stimuli. However, given an echoic repertoire, the sequence in which the training of listener behavior, tacts, and intraverbals is programmed should not necessarily affect the ease with which all three relations are acquired.

Despite this notion of functional independence, everyday experience suggests that typically developing children and adults frequently emit speaker and listener responses that have not previously been reinforced. It is also likely that many clinicians implementing EIBI programs for children with autism have observed the emergence of untrained speaker or listener relations, especially following the acquisition of some rudimentary verbal skills. Skinner (1957) acknowledged that it sometimes appears as if following the acquisition of one verbal operant, another emerges without any prior history of reinforcement. Skinner assumed that such apparently spontaneous transfer could be accounted for in terms of known behavioral principles, such as multiple stimulus control and the unobserved occurrence of collateral responses that are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reinforced at the time of acquisition. In this context, he described in detail, for example, various circumstances that might give rise to transfer of control between mand and tact contingencies (pp. 188-189). With respect to speaker and listener behavior, Skinner did not elaborate on the specific variables that might lead a listener to respond appropriately to an instruction such as /Point to the cat./, having previously learned only to tact a cat, or vice versa. However, other authors have more recently offered conceptual analyses of such performances (e.g., D. Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000; Horne & Lowe, 1996; Lowenkron, 1998), some of which will be discussed in later sections.

If some degree of interdependence may exist between speaker and listener relations, then the training of certain relations may either generate untrained relations or facilitate their acquisition. This raises numerous questions. For example, does the training of listener relations facilitate acquisition of tacts and intraverbals, as the sequencing of programs in EIBI curricula appears to suggest? Would initial training of tacts or intraverbals be equally likely to facilitate or generate listener relations? Answers to these questions appear to be among the empirical evidence needed in order to determine the most efficient sequencing of programs. Before reviewing the relevant literature, however, it is necessary to consider again the distinction between selection-based and topography-based behavior.

Selection-based versus Topography-based Relations and Stimulus Equivalence

> If a pigeon's pecks are reinforced in the presence of green but not red, we wouldn't be likely to consider the possiblity of the pigeon's greening in the presence of pecks. (Catania, 1998, p. 151)

Language training programs that employ selection-based communication systems may be able to benefit from a large body of behavior-analytic research that has

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

been conducted in the area of *stimulus equivalence* (for general reviews, see Green & Saunders, 1998; Sidman, 1994). This research has revealed that the training of certain relations may result in the predictable emergence of other relations in humans. Research on stimulus equivalence has typically been conducted using arbitrary match-to-sample procedures. In this type of preparation, the experimental task consists of selecting (e.g., by touching) a comparison stimulus from an array, following the presentation of a given sample stimulus that bears no physical similarity to the correct comparison (Green & Saunders, 1998). Thus, conditional discriminations are trained and tested. A typical finding is that following training of two baseline relations, for example, selecting stimulus B given stimulus A, and stimulus C given stimulus B, a number of other relations emerge without further training. The emergence of these relations is said to indicate that training has established the A, B, and C stimuli as members of a common equivalence class, which has been defined as a stimulus class that exhibits the properties of reflexivity, symmetry, and transitivity (Sidman & Tailby, 1982). Reflexivity is demonstrated when the participant selects each stimulus in the presence of an identical stimulus; that is, A given A, B given B, and so forth. Symmetry is demonstrated when the participant can, without additional training, select A given B and B given C, and transitivity is demonstrated when the participant can select C given A, even though the two stimuli have never before been presented together. The selection of A in the presence of C can emerge only if the properties of reflexivity, symmetry, and transitivity are met, and the emergence of this relation alone is thus said to suffice to demonstrate the establishment of an equivalence class (Sidman, 1994; Sidman & Tailby, 1982).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The sample and comparison stimuli, although frequently visual, may be presented in other sensory modalities as well (Dube, Green, & Serna, 1993; L. J. Hayes, Tilley, & Hayes, 1988) and they may be either verbal stimuli such as spoken or written words, or nonverbal stimuli such as objects or pictures. When the sample stimulus is verbal, the selection of a nonverbal comparison stimulus is an instance of listener behavior, according to Skinner's (1957) analysis. When the sample stimulus is nonverbal, the selection of a verbal comparison stimulus may be thought of as a selection-based tact, and if both sample and comparison stimuli are verbal, the response may be considered a selection-based intraverbal.

The emergence of equivalence responding following an appropriate experimental history has been observed reliably in human adults, and has also been observed in individuals with limited verbal repertoires due to developmental disabilities (for a review, see O'Donnell & Saunders, 2000) as well as in children of preschool age (e.g., Devany, Hayes, & Nelson, 1986; Smeets & Barnes-Holmes, 2005). With regard to the emergence of untrained speaker and listener relations, it may be noted, for example, that selection-based tacts, such as selecting a printed word given a picture of an object as a sample stimulus, tend to emerge reliably as a result of reinforcement of the symmetrical listener relations, and vice versa (e.g., Sidman, Cresson, & Wilson-Morris, 1985). In general, there is no reason to suppose that substituting, for example, written verbal stimuli for other arbitrary visual stimuli that have employed in the literature would yield different results. Therefore, to the extent that equivalence performances are observed in individuals with limited verbal repertoires, the reinforcement of a selection-based intraverbal relation should generate an opposite selection-based intraverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

relation as an instance of symmetry; the reinforcement of two listener relations involving the same nonverbal but different verbal stimuli should generate selection-based intraverbals as instances of transitivity; and so forth. As a result, the extensive research programs conducted in this area may be able to directly inform the structure of language training programs in which the verbal operants trained are selection-based.

The same may not be true for programs that teach topography-based verbal responding. At this point, it may be worth noting that a number of different theoretical perspectives exist on the origin of the emergent performances observed in stimulus equivalence research, as well as the relationship of such performances with verbal behavior (for reviews and discussion, see Clayton & Hayes, 1999; Zentall, Galizio, & Critchfield, 2002). One theory (S. C. Hayes, Barnes-Holmes, & Roche, 2001) views responding in accordance with symmetry and transitivity on arbitrary match-to sample tasks as instances of *arbitrarily applicable relational responding*. Various types of arbitrarily applicable relational responding are considered to be generalized higher-order operants that emerge from a particular history of reinforcement of *mutually entailed* (a more general term for symmetrical) and *combinatorially entailed* (a more general term for transitive) relations in the presence of relevant *contextual cues* (e.g., instructions or other verbal or nonverbal stimuli). Further, arbitrarily applicable relational responding is considered to be a prerequisite to verbal behavior, or more precisely, a prerequisite for untrained speaker or listener relations emerging as a result of the establishment of others (D. Barnes-Holmes et al., 2000). From this point of view, according to D. Barnes-Holmes et al., topography-based tacts and listener relations involving the same nonverbal stimulus are seen as mutually entailed relations. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

emergence of one as a result of the establishment of the other is said to exemplify the same process by which other mutually entailed relations are derived, and thus reflect the same type of reinforcement history. Other mutually or combinatorially entailed relations are seen as capable of arising in a similar manner. An implication of this view therefore appears to be that findings from the stimulus equivalence literature may be directly applicable to language-training programs, and the distinction between selection-based and topography-based responding may not be of fundamental importance.

A different point of view is that the types of emergent performances seen in research on stimulus equivalence are in fact mediated by topography-based verbal responding that may occur overtly or covertly (e.g., Horne & Lowe, 1996; Lowenkron, 1998). In other words, topography-based verbal relations are considered prerequisite for equivalence responding. From this point of view, it appears that the emergence of untrained topography-based relations from the training of other relations warrants its own research program. A third point of view, held by Sidman (1994) among others, is that emergent performances on stimulus equivalence tests are probably neither learned nor dependent on the occurrence of verbal behavior. Although these performances are typically considered in some way related to language and cognitive skills, language is not necessarily viewed as an instance of such performances. Sidman (1994) has stated, despite earlier suggestions to the contrary (Sidman & Tailby, 1982), that a vocal speaker relation, such as tacting a cat, cannot in fact be considered symmetrical to the listener relation in which a cat is selected from an array of stimuli in the presence of the verbal stimulus /cat/. Rather, symmetry of the latter relation would be demonstrated by the ability to select the stimulus /cat/ from a limited array of vocally produced stimuli.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Producing the vocal response requires an additional topography that is not a part of the trained listener relation. According to Sidman, "[w]e cannot, therefore, take accurate naming to imply accurate auditory-visual matching. Nor can we automatically take accurate auditory-visual matching to imply accurate naming." (1994, p. 228). Similar views regarding stimulus-response reversibility have been presented by others (e.g., Hall & Chase, 1991), implying that findings from the stimulus equivalence literature may not be directly applicable to the training of topography-based relations.

To date, the theoretical debate on the origin of derived stimulus relations has not been resolved. Neither does the stimulus equivalence literature at present indicate whether the training of topography-based relations in fact yields similar or different results from those yielded by the training of selection-based relations. In research on stimulus equivalence, participants with limited verbal repertoires have sometimes been tested for the emergence of vocal tacts and other vocal speaker relations following training of selection-based listener relations. While speaker relations have frequently emerged (e.g., Brady & McLean, 2000; Mackay, 1985; Mackay & Ratti, 1990), they have also failed to do so (e.g., Sidman & Cresson, 1973; Sidman et al., 1974; Sidman, Wilson-Morris, & Kirk, 1986). On the other hand, topography-based relations have not typically been *trained* in research on stimulus equivalence, as such training is incompatible with the typical match-to-sample training paradigm. When topography-based relations have been trained, the primary purpose has often been to assess whether stimulus control over the topography-based response transfers across members of an equivalence class, and thus the emergence of listener relations has not been assessed (e.g., Barnes, Brown, Smeets, & Roche, 1995; Bones et al., 2001). The stimulus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

equivalence literature to date may thus provide little information on whether speaker and listener training may be differentially effective in terms of generating both listener and speaker relations, when the speaker behavior in question is topography-based. A few recent studies exist within this literature, however, in which topography-based responses to nonverbal stimuli have been both trained and tested along with corresponding listener relations (Y. Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001a, 2001b; Lipkens, Hayes, & Hayes, 1993), or in which identical methods have been employed across studies to investigate the effects of tact and listener training (Horne, Hughes, & Lowe, 2006; Horne, Lowe, & Randle, 2004; Lowe, Horne, Harris, & Randle, 2002; Lowe, Horne, & Hughes, 2005). Those studies are reviewed in the following section along with a number of studies from outside of the stimulus equivalence literature, in which topography-based tacts and listener relations have also been trained and tested. This is followed by a review of studies that have assessed the effects of training intraverbals on listener relations or vice versa.

Listener Behavior and the Tact

Listed in Tables 1 through 3 are 23 experimental studies that have assessed the effects of tact training on the emergence of listener relations, as well as the effects of listener training on tact emergence, in participants with limited verbal abilities. Studies were selected for review if the participants were children or adults with diagnoses of mental retardation or autism, or any children under the age of six. One study was included, however, in which at least one of four participants had reached age six (Williams & McReynolds, 1975). In general, only studies were reviewed in which the training and testing of both tacts and listener relations was reported within the same

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

article. However, four studies were included in which the training of only one of these repertoires was the focus of study, but in which the same researchers employed identical procedures across studies to investigate the effects of either tact or listener training with similar participants (Horne et al., 2004, 2006; Lowe et al., 2002, 2005).

Table 1
Number of Participants and Experimental Arrangement in Studies that Have Evaluated the Effects of Tact and Listener Training on Emergent Tact and Listener Repertoires.

Study	N	Age in years	*n* tested for tacts	*n* tested for listener relations	Within / between subjects	Training arrangement	Testing arrange-ment
MR, other than autism:							
Cuvo & Riva (1980)	10	Mean: 16.2	5	5	Between	n/a	Post
Goldstein et al. (1987)	3	7-18	3	3	Within	Sequential	Repeated
Guess & Baer (1973; Exp. 1)	4	11-21	4	4	Within	Simultaneous	Repeated
Holdgrafer (1981)	2	14-15	1	1	Between	n/a	Post
Keller & Bucher (1979)	6	Children	6	6	Within	Sequential	Repeated
Lee (1981; Exp. 1 and 2)	2	9-10	2	2	Within	Sequential	Repeated
Miller et al. (1977)	14	Mean: 12.7	7	7	Between	n/a	Post
Smeets (1978)	1	14	1	1	Within	Simultaneous	Repeated
Smeets & Striefel (1976)	1	14	1	1	Within	Simultaneous	Repeated
Total	42[a]		29[a]	29[a]			
Autism:							
Eikeseth & Jahr (2001)	4	4-7	2	3	Within[b]	Simultaneous	Post
Watters et al. (1981)	4	10-16	4	4	Within	Simultaneous	Repeated
Wynn & Smith (2003)	6	3-6	6	6	Within	Sequential	Repeated
Total	14		12	13			

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1–Continued

Study	N	Age in years	*n* tested for tacts	*n* tested for listener relations	Within / between subjects	Training arrangement	Testing arrangement
Other children under 6:							
Y. Barnes-Holmes et al. (2001a)[a]	16	4-5	12	4	Between	n/a	Repeated
Y. Barnes-Holmes et al. (2001b)[a]	16	4-5	12	4	Between	n/a	Repeated
Connell (1986)	6	2-3	3	3	Between	n/a	Repeated
Connell & McReynolds (1981)	6	5	6	6	Within	Sequential	Post
Cuvo & Riva (1980)	10	Mean: 4.4	5	5	Between	n/a	Post
Eikeseth & Jahr (2001)	3	3	3	3	Within	Simultaneous	Post
Holdgrafer & McReynolds (1975)	4	3	4	4	Within	Simultaneous	Repeated
Horne et al. (2004)	9[c]	1-4	9	0	n/a	n/a	Post
Horne et al. (2006)	14	1-4	14	0	n/a	n/a	Post
Lipkens et al. (1993; Exp. 1, 2, and 4)	1	1-2	1	1	Within	Sequential	Repeated
Lowe et al. (2002, Exp. 2)	3	3-4	0	3	n/a	n/a	Post
Lowe et al. (2005)	9[d]	1-3	0	9	n/a	n/a	Post
Williams & McReynolds (1975)	4	5-6	4	4	Within	Sequential	Post
Total	101		73	46			
Total	157		114	88			

[a] The same adolescent participated in Smeets (1978) and Smeets and Striefel (1976).

[b] All participants received both tact and listener training, but some participants did not acquire the directly trained repertoires and were therefore not tested for the emergence of the other.

[c] 20 children participated but only nine reached the testing phase.

[d] 10 children participated but one did not reach the testing phase.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2
Training Tasks, Response Topographies, and Stimuli Used in Studies that Have Evaluated the Effects of Tact and Listener Training on Emergent Tact and Listener Repertoires.

Study	Training task	Tact modality	Tact topography	Listener topography	Nonverbal stimuli	Size of listener array
MR, other than autism:						
Cuvo & Riva (1980)	Names	Vocal	Conventional	Point	3-D conv.	5
Goldstein et al. (1987)	Sentences	Vocal	Conventional	Object placement	3-D conv.	6-8
Guess & Baer (1973; Exp. 1)	Plural suffix	Vocal	Conventional	Point	3-D conv.	2
Holdgrafer (1981)	Plural suffix	Vocal	Conventional	Point	2-D conv.	Not stated
Keller & Bucher (1979)	Names	Vocal	Conventional	Not stated	2-D, conv..	Not stated
Lee (1981; Exp. 1 and 2)	Sentences	Vocal	Conventional	Object placement	3-D conv.	2
Miller et al. (1977)	Names	Vocal	Conventional	Point	3-D conv.	Not stated
Smeets (1978)	Plural suffix	Motor	Conventional	Point	2-D conv.	2
Smeets & Striefel (1976)	Names	Motor	Conventional	Point, touch	2-D conv.	8
Autism:						
Eikeseth & Jahr (2001)	Names	Motor	Conventional	Not stated	2-D conv.	3
Watters et al. (1981)	Names	Motor	Conventional	Give	2-D conv.	5
Wynn & Smith (2003)	Names/ attributes	Vocal	Conventional	Touch	3-D conv.	2
Other children under six:						
Y. Barnes-Holmes et al. (2001a)[a]	Names	Motor	Nonsense	Not stated	3-D conv.	2
Y. Barnes-Holmes et al. (2001b)[a]	Names	Motor	Nonsense	Not stated	3-D conv.	2
Connell (1986)	Sentences	Vocal	Conventional	Point	2-D conv.	2
Connell & McReynolds (1981)	Names	Vocal	Nonsense	Touch	2-D arb.	3
Cuvo & Riva (1980)	Names	Vocal	Conventional	Point	3-D conv.	5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2–Continued

Study	Training task	Tact modality	Tact topography	Listener topography	Nonverbal stimuli	Size of listener array
Eikeseth & Jahr (2001)	Names	Motor	Conventional	Not stated	2-D conv.	3
Holdgrafer & McReynolds (1975)	Plural suffix	Vocal	Conventional	Point	2-D conv.	Not stated
Horne et al. (2004)	Common names	Vocal	Nonsense	Give	3-D arb.	2
Horne et al. (2006)	Common names	Vocal	Nonsense	Give	3-D arb.	2
Lipkens et al. (1993; Exp. 1, 2 and 4)	Names	Vocal	Nonsense	Touch	2-D arb.	2
Lowe et al. (2002; Exp. 2)	Common names	Vocal	Nonsense	Give	3-D arb.	2
Lowe et al. (2005)	Common names	Vocal	Nonsense	Give	3-D arb.	2
Williams & McReynolds (1975)	Names	Vocal	Nonsense	Place in box	2-D arb.	4

Table 3

A Summary of Results from Studies that Have Evaluated the Effects of Tact and Listener Training on Emergent Tact and Listener Repertoires.

Study	Overall performance		Terminal performance			
	Pass / fail tact tests	Pass / fail listener tests	Pass / fail tact tests	Pass / fail listener tests	Functional independence	Functional inter-dependence
MR, other than autism:						
Cuvo & Riva (1980)	n/a (mean: 75% correct)	n/a (mean: 96% correct)			n/a	n/a
Goldstein et al. (1987)	1 / 1	2 / 0	1 / 2	2 / 1	0	1
Guess & Baer	0 / 2	0 / 1	1 / 3	1 / 3	1	0
(1973; Exp. 1)						
Holdgrafer (1981)	0 / 2	0 / 2			n/a	n/a
Keller & Bucher (1979)	0 / 6	5 / 1			Not known	0
Lee (1981; Exp. 1 and 2)	0 / 2	1 / 0	0 / 2	2 / 0	0	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3–Continued

Study	Overall performance		Terminal performance			
Miller et al. (1977)	n/a (mean: not stated)	n/a (mean: 99% correct)			n/a	n/a
	Pass / fail tact tests	Pass / fail listener tests	Pass / fail tact tests	Pass / fail listener tests	Functional independence	Functional inter-dependence
Smeets (1978)	0 / 0	1 / 0	1 / 0	1 / 0	0	0
Smeets & Striefel (1976)	0 / 0	1 / 0	1 / 0	1 / 0	0	0
Total	1 / 13	9[a] / 4	3[a] / 7	6[a] / 4	1	1
Autism:						
Eikeseth & Jahr (2001)	0 / 2	1 / 2			0	0
Watters et al. (1981)	0 / 4	2 / 2			0	0
Wynn & Smith (2003)	0 / 3	1 / 1	1 / 5	2 / 4	0	0
Total	0 / 9	4 / 5	1 / 5	2 / 4	0	0
Other children under six:						
Y. Barnes-Holmes et al. (2001[a])[a]	2 / 0	1 / 0	12 / 0	4 / 0	n/a	n/a
Y. Barnes-Holmes et al. (2001b)[a]	4 / 0	0 / 0	12 / 0	4 / 0	n/a	n/a
Connell (1986)	0 / 3	0 / 3	0 / 3	0 / 3	n/a	n/a
Connell & McReynolds (1981)	1 / 5	5 / 1			1	1
Cuvo & Riva (1980)	n/a (mean: 75% correct)	n/a (mean: 96% correct)			n/a	n/a
Eikeseth & Jahr (2001)	1 / 2	1 / 2			0	0
Holdgrafer & McReynolds (1975)	0 / 0	1 / 0	1 / 3	1 / 3	0	0
Horne et al. (2004)	2 / 7	n/a	3 / 6	n/a	n/a	n/a
Horne et al. (2006)	10 / 4	n/a			n/a	n/a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3–Continued

Study	Overall performance		Terminal performance			
Lipkens et al. (1993; Exp. 1, 2 and 4)	0 / 0	0 / 0	1 / 0	1 / 0	0	0
Lowe et al. (2002; Exp. 2)	n/a	3 / 0			n/a	n/a
	Pass / fail tact tests	Pass / fail listener tests	Pass / fail tact tests	Pass / fail listener tests	Functional independence	Functional inter-dependence
Lowe et al. (2005)	n/a	9 / 0			n/a	n/a
Williams & McReynolds (1975)	0 / 4	3 / 1			1	0
Total	20 / 25	23 / 7	29 / 12	10 / 6	2	1
Total (all studies)	21 / 49	36 / 16	34 / 24	19 / 14	3	2
Percentage of *n* tested	18.4% / 43.0%	40.1% / 18.2%	58.8% / 41.2%	57.6% / 42.4%		

[a] The participant in Smeets (1978) and Smeets and Striefel (1976) is only counted once.

The 23 studies do not represent a systematic line of research, which reflects their origins in different fields of study; such as the experimental analysis of behavior, applied behavior analysis, speech and language pathology, and applied psycholinguistics. The earliest study was conducted by Guess and Baer (1973) for the purpose of evaluating the effects of the two types of training for teaching the plural form of nouns. This study was a systematic replication and extension of a study by Guess (1969), who investigated the effects of listener training only. A number of early studies appear to have been directly informed by Guess and Baer's study, and some are identified as replications or extensions of it (Holdgrafer, 1981; Holdgrafer &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

McReynolds, 1975; Lee, 1981; Smeets, 1978). Other studies arose out of different considerations, and not all of them were conducted explicitly for the purpose of comparing the effects of tact and listener training, but rather produced relevant data indirectly. For example, Y. Barnes-Holmes et al. (2001a, 2001b) conducted a series of seven experiments, the purpose of which was to investigate the effects of various procedures on the extent to which the training of either relation would generate the other. Similarly, Lipkens et al. (1993) conducted a longitudinal study of the acquisition of speaker and listener relations by a typically developing infant, and although its purpose was to evaluate the emergence of untrained relations, it was not specifically concerned with comparing the effects of different types of training. Horne, Lowe, and colleagues (Horne et al., 2004, 2006; Lowe et al., 2002, 2005) examined the effects of tact and listener training on stimulus sorting performances. The purpose was to investigate the extent to which the emergence of such performances, as well as the transfer of new discriminative functions (Horne et al., 2006; Lowe et al., 2005), would correlate with the presence of *name relations*. According to these authors (Horne & Lowe, 1996), a name relation is demonstrated when a child exhibits appropriate listener behavior following tact training or vice versa. Thus, the participants in these studies were tested for emergent tacts or listener relations following either tact or listener training. The purpose of Eikeseth and Jahr's (2001) study was to compare the effects of a topography-based sign training program with the effects of a program designed to teach selection-based sight-reading skills; however, the evaluation incorporated a measure of tact emergence following listener training and vice versa. Goldstein et al. (1987) were concerned with evaluating the effects of a procedure termed matrix training

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on recombinative generalization; that is, the extent to which training with carefully selected pairs of stimuli resulted in generalization to novel stimulus combinations. However, they additionally assessed the effects of listener training on tact emergence and vice versa. Finally, the purpose of Watters et al. (1981) was to compare the effects of two training procedures termed expressive and receptive simultaneous communication training. In expressive simultaneous communication, the learner is trained to emit a sign in the presence of a nonverbal stimulus as well as its spoken name. The resulting relation could either be a tact or an intraverbal, or an instance of multiple control by the nonverbal and verbal stimulus over signing. In receptive simultaneous communication training, the learner selects a nonverbal stimulus when simultaneously presented with a sign and a spoken word, as a result of which the selection response may be under the control of one or both of the verbal stimuli. Although it was thereby not clear whether the trained and tested speaker relations were tacts or intraverbals, stimulus control tests over trained and tested listener repertoires indicated that they were primarily under the control of signs rather than spoken words, and thus this study is included as an example of research on listener relations and tacts.

It should be noted that only a small minority of the studies reviewed made use of Skinner's (1957) verbal operant taxonomy; in fact, only Horne et al. (2004, 2006) and Lowe et al. (2002, 2005) did so. Other researchers described their target repertoires as receptive and expressive or productive language (Guess & Baer, 1973; Eikeseth & Jahr, 2001; Goldstein et al., 1987; Keller & Bucher, 1979; Smeets, 1978; Smeets & Striefel, 1976; Watters et al., 1981; Wynn & Smith, 2003) , comprehension or discrimination and production (Connell, 1986; Connell & McReynolds, 1981; Cuvo & Riva, 1980;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Holdgrafer, 1981; Holdgrafer & McReynolds, 1975; Miller, Cuvo, & Borakove, 1977; Williams & McReynolds, 1975), verbal and nonverbal behavior (Lee, 1981), name-object and object-name relations (Lipkens et al., 1993), and action-object and object-action relations (Y. Barnes-Holmes et al., 2001a, 2001b; these researchers, however, did use Skinner's taxonomy when describing pretraining conditions involving vocal rather than motor topographies). In the present review, for ease of comparison, vocal or motor responses to nonverbal stimuli will be referred to as tacts, whereas the terms listener behavior and listener relations will be used to denote selection of nonverbal stimuli in the presence of a verbal stimulus, as well as such stimulus-selection combined with simple mand compliance.

A total of 157 individuals participated in the studies; 114 were tested for tact emergence following listener training, and 88 were tested for the emergence of listener behavior following tact training. Tables 1 through 3 describe details of each study. The number of participants who received each type of training is described in Table 1, along with basic information on method. Table 2 contains information on stimuli and response topographies trained and tested, and results are summarized in Table. 3. Two studies (Cuvo & Riva, 1980; Eikeseth & Jahr, 2001) employed participants both with and without disabilities, and each participant group is analyzed separately in the tables. Typically developing adults, in addition to children, participated in one study (Connell & McReynolds, 1981), but information and data on the adults are not included in the tables. In Guess and Baer's (1973) study, only Experiment 1 is included, as no untrained relations were tested in Experiment 2. In Lee's (1981) study, only Experiments 1 and 2 are included, as Experiment 3 employed different procedures than

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

those used in the previous experiments to investigate the effects of tact training only, with new participants. Experiment 3 in Lipkens et al. (1993) and Experiment 1 in Lowe et al. (2002) are omitted because no listener relations were either trained or tested in those experiments. Finally, only data on children who participated in Treatment B in Connell and McReynolds (1981) are reported, because Treatment A focused on transfer of function across arbitrarily related stimuli more than on the emergence of untrained tact and listener relations.

Experimental arrangement. Methodologically, the studies listed in Tables 1 through 3 were very heterogeneous. Five studies (Guess & Baer, 1973; Keller & Bucher, 1979; Lee, 1981; Smeets, 1978; Smeets & Striefel, 1976) employed experimental single-case designs to directly compare the effects of tact and listener training. Several other small-N studies (Connell & McReynolds, 1981; Eikeseth & Jahr, 2001; Goldstein et al., 1987; Holdgrafer & McReynolds, 1975; Lipkens et al., 1993; Watters et al., 1981; Wynn & Smith, 2003; Williams & McReynolds, 1975) evaluated the effects of the two types of training within subjects, but employed designs that strictly speaking may not have permitted a comparison of the two procedures. In some cases (Goldstein et al.; Lipkens et al.), this was because the study was not designed for the purpose of comparing them. Two studies (Cuvo & Riva, 1980; Miller et al., 1977) employed between-subjects designs to compare the effects of tact and listener training. In the remaining studies (Y. Barnes-Holmes et al., 2001a, 2001b; Connell, 1986; Holdgrafer, 1981; Horne et al.,. 2004, 2006; Lowe et al., 2002, 2005), different participants received tact and listener training, but no between-subjects design was employed to compare the two types of training, in most cases because this was not the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

purpose of the study (Y. Barnes-Holmes et al., 2001a, 2001b; Horne et al., 2004, 2006, Lowe et al., 2002, 2005).

As with experimental design, training and testing arrangement varied widely across studies. In the 14 studies that evaluated the effects of tact or listener training within subjects, participants received the two types of training either simultaneously or sequentially (see Table 1). Seven studies employed simultaneous training, in which participants received tact and listener training on different stimulus sets. In Eikeseth and Jahr's (2001) study, simultaneous training was arranged in such a way that tact and listener training were alternated across sessions, and post-tests of untrained tact and listener relations were administered following the completion of both training conditions. In the remaining six studies that employed simultaneous training (Guess and Baer, 1973; Holdgrafer & McReynolds, 1975; Smeets, 1978; Smeets & Striefel, 1976; Watters et al., 1981), training on exemplars or exemplar pairs from different stimulus sets was alternated such that following mastery on one exemplar or pair in tact training, listener training was employed for the next exemplar or pair, and vice versa. In all six studies, untrained tact and listener relations were tested repeatedly, following training on particular exemplars or exemplar pairs. However, Watters et al. reported only aggregate data on the outcome of those tests.

In the remaining seven studies in which the same participants received tact and listener training, the two training conditions were ordered sequentially. For example, participants would receive tact training on an entire stimulus set, followed by listener training on another. In Lee's (1981) study, tact and listener training conditions were alternated multiple times, and in Keller and Bucher's (1979) study, all participants were

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

exposed to each condition twice. In two additional studies (Goldstein et al., 1987; Wynn & Smith, 2003), each participant was exposed to at least one of the training conditions more than once, and the order of conditions varied across participants. In those four studies, the untrained repertoire was tested repeatedly during (Lee; Keller & Bucher; Wynn & Smith) or following (Goldstein et al.) training, but Keller and Bucher reported only aggregate data. In two studies (Connell & McReynolds, 1981; Williams & McReynolds, 1975), each participant was exposed to each training condition once, with the order of conditions counterbalanced across participants, and a post-test was administered at the end of each condition. Finally, in Lipkens et al. (1993), training was arranged in such a way that in Experiment 1, tact training was followed by listener training, tact training only was investigated in Experiment 2, and listener training only in Experiment 4. In that study, the untrained repertoire was in all cases tested repeatedly throughout training.

In the two group-design studies (Cuvo & Riva, 1980; Miller et al., 1977), each group received either tact or listener training, and pre- and post-tests were used to assess the emergence of the untrained repertoire. Of the remaining eight studies in which different participants received tact and listener training, the untrained repertoire was measured repeatedly during training in one study (Connell, 1986), and in two studies (Y. Barnes-Holmes et al.; 2001a, 2002b), it was measured repeatedly following training on different stimulus sets. The remaining five studies reported only post-training performance on a single test (Holdgrafer, 1981; Horne et al., 2004, 2006; Lowe et al., 2002, 2005), except that Horne et al. (2004) reported repeated instances of testing for three participants, following additional training.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Other differences in experimental design may be noted. For example, several studies employed contingency reversal procedures in which, following initial training of tacts and/or listener relations, one or more training conditions followed in which verbal stimuli and response topographies were reassigned to nonverbal stimuli (Connell & McReynolds, 1981; Holdgrafer, 1981; Holdgrafer & McReynolds, 1975; Lee, 1981; Smeets, 1978; Williams & McReynolds, 1975). As another example, several studies (Y. Barnes-Holmes et al., 2001a, 2001b; Connell, 1986; Cuvo & Riva, 1980; Miller et al., 1977; Wynn & Smith, 2003) demonstrated experimental control over the untrained repertoire by training it directly if it did not emerge following training of the other repertoire. In one study, this was the focus of a separate experiment (Guess & Baer, 1973).

Response topographies trained and tested. The "Training Task" column in Table 2 describes the focus of each study. "Names" is listed as the training task in 12 studies. In those studies, one verbal stimulus and one verbal response topography corresponded to each nonverbal stimulus. When conventional topographies were trained, the names in question were typically nouns, but this was not necessarily the case. For example, the training set in Smeets and Striefel (1976) included "green"; however, only one nonverbal stimulus corresponded to this topography. In Wynn and Smith's (2003) study, names were the training task for two participants, whereas the other four were trained on "Attributes"; that is, abstract stimulus properties shared by a number of training exemplars, such as hot/cold, and long/short. In four studies, the task was "Common names", in which each verbal stimulus and response topography corresponded to multiple nonverbal stimuli that did not share physical properties.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"Plural suffix" was the training task in four studies. In those studies, tact training involved training the participants to emit different response topographies based on the number of nonverbal stimuli present. In listener training, the participants learned to respond to different verbal stimuli by selecting a stimulus configuration that contained the appropriate number of items. Conventional plurality and singularity was trained in three of those studies (Guess & Baer, 1973; Holdgrafer, 1981; Smeets, 1978), whereas Holdgrafer and McReynolds (1975) trained a distinction between "many" and "a few" using an artificial suffix. Finally, in three studies, the training task was sentences. In Connell's (1986) study, the sentences contained an actor, an action, and an object (e.g., "worm kiss frog"). The participants were previously able to tact all actors, objects and actions when presented separately. In Lee's (1981) study, the sentences described the location of a target object relative to another (e.g., "on the right of the cup"), and the participants were pretrained to tact all objects. In the study by Goldstein et al. (1987), the sentences specified an action that remained constant across exemplars ("put"), two objects, and their relative locations (e.g., "put the balloon on the bed") and the participants were able to tact some but not all objects prior to the study.

As shown in Table 2, the trained and tested tact topographies differed somewhat from one study to another. Vocal tacts were trained and tested in a majority of the studies. In some cases, the topographies were conventional ones from the participants' native languages, and in others they were nonsense topographies. Four studies employed conventional or modified sign language topographies (Eikeseth & Jahr, 2001; Smeets, 1978; Smeets & Striefel, 1976; Watters et al., 1981) and two studies (Y. Barnes-Holmes et al., 2001a, 2001b) employed other motor topographies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the listener training and testing conditions, the target repertoires in all except two studies consisted of pointing to or otherwise selecting a nonverbal stimulus from an array of stimuli. In the remaining two studies (Goldstein et al., 1987; Lee, 1981), listener responding consisted of placing one object in a specific location relative to another object, and thus required three conditional discriminations. In the Goldstein et al. study, a number of objects were available to select from, but only two were available in Lee's study.

In most cases, the nonverbal stimuli used for listener training were the same as those used for testing tacts in that study, and the stimuli used for testing listener relations were the same as those used for training tacts. Three studies, however (Connell, 1986; Goldstein, 1987; Lee, 1981), did not assess untrained relations directly with the training stimuli, but assessed generalization of both trained and untrained repertoires to novel exemplars. Three additional studies (Guess & Baer, 1973; Holdgrafer & McReynolds, 1975; Smeets, 1978) incorporated generalization tests for the trained repertoire, and Holdgrafer and McReynolds probed generalization in the untrained repertoire as well. Finally, in addition to testing with the training stimuli, Connell and McReynolds (1981) assessed transfer of discriminative control over the untrained repertoire to stimuli that the children had previously learned to match to the training stimuli in an arbitrary visual match-to-sample task; that is, they assessed derived tacts and listener relations.

The nonverbal stimuli were visual in all cases. As shown in Table 2, the stimuli could be either two-dimensional or three-dimensional, and they could represent either familiar items, or consist of arbitrary shapes or drawings. The size of the array used for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

testing differed from one study to another (see Table 2). Arrays included from 2 stimulus items up to 8 items (Smeets & Striefel, 1976); however, the most common array size was two items, employed in 13 studies.

Analysis of results. In Table 3, the results of each study are analyzed in several ways. The "Overall performance" columns show participants' overall performance during the study, or their performance on post-tests at the completion of the study or each of its conditions. The first column shows (a) the number of participants who consistently passed tact tests following listener training, and (b) the number of participants who consistently failed tact tests following listener training. The second column shows (a) the number of participants who consistently passed listener tests following tact training, and (b) the number of participants who consistently failed listener tests following tact training. Participants who did not either consistently pass or fail are not included in these columns. In order to determine whether a participant passed or failed particular tests, the passing criterion used by the researchers was employed. If the researchers did not describe a passing criterion, 90% correct performance was considered passing. A participant was considered to have passed "consistently" if he or she passed at least 90% of all instances of testing when repeated measures were employed, but if a post-test was used or only aggregate measures reported, the participant passed that test.

Studies in which participants were repeatedly exposed to each training and testing condition allowed for an additional analysis of performance following the last instance of each type of training. Because in a few studies there was evidence of improvement across repeated instances of training and testing, these terminal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

performance data are reported in Table 3 in the same manner as overall performance data. Following the last instance of training, each participant necessarily either passed or failed the subsequent test, as a result of which the sum of passing and failing participants in those columns equals the number of participants who were tested for the emergence of a particular relation.

Under the “Complete functional independence” heading are listed the number of participants in each study for whom *no increase* in correct responding was observed on tact tests following listener training and listener tests following tact training. Participants who showed some consistent evidence of listener or tact emergence, but did not meet the passing criterion for either repertoire, are not included under this heading. The “Complete functional interdependence” column, by contrast, lists the number of participants who consistently passed all tests following both types of training.

Results for participants with mental retardation. Children and/or young adults with mental retardation, but without diagnoses of autism, participated in nine studies. As shown in Table 1, an equal number of participants in those studies received tacts and listener training. As Table 3 shows, Guess and Baer (1973) demonstrated functional independence of tacts and listener relations with one of their participants. In addition, neither of Holdgrafer’s (1981) two participants acquired the untrained repertoire. For all other participants in the small-N studies, some interdependence between the two repertoires was observed, as many participants acquired at least one of the untrained repertoires partially or fully. One participant fully acquired both of the untrained repertoires if overall performance is considered (Goldstein, 1987). Two additional participants, one of whom participated in two studies, ultimately passed both tact and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

listener tests at the end of the study, following multiple instances of training and testing of both types of relations (Guess & Baer; Smeets, 1978; Smeets & Striefel, 1976). One of those participants, however (Smeets; Smeets & Striefel), overall appeared to acquire untrained listener relations more readily than tacts. Of the remaining 14 participants who participated in within-subject evaluations, there was similarly evidence for 10 participants that tact training more readily generated listener relations than vice versa (Guess & Baer; Goldstein et al.; Keller & Bucher, 1979; Lee, 1981), even though not all of these participants consistently passed all listener tests. The opposite pattern was never reported. Overall, 9 out of 17 participants in within-subject evaluations consistently passed, and 4 consistently failed, listener tests following tact training. By contrast, only 1 out of 17 consistently passed tact tests following listener training, whereas 13 consistently failed.

One of the two group studies (Miller et al., 1977) also indicated a greater efficiency of tact than listener training in terms of generating untrained relations, whereas the difference observed in the other (Cuvo & Riva, 1980) did not reach statistical significance. Both of the group studies additionally reported that tact training required less time and/or fewer trials to criterion than listener training, as did one of the within-subject design studies (Smeets, 1978).

The two studies that tested generalization to novel exemplars within the trained repertoire (Guess & Baer, 1973; Smeets, 1978) both reported high levels of generalization as a result of both tact and listener training. In Guess and Baer's study, however, there was a tendency for three participants to perform better on listener

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

generalization tests following listener training than on tact generalization tests following tact training.

Results for participants with autism. Children with autism participated in three studies, all within-subject evaluations. A total of 12 children were tested for tact emergence following listener training and 13 for the emergence of listener relations following tact training (see Table 1). The results overall closely mirrored those obtained with participants with other developmental disabilities. As shown in Table 3, complete functional independence of the two repertoires was never observed as defined here; however, one participant in Eikeseth and Jahr's (2001) study showed no emergence of listener relations following tact training, and responded with only 10% accuracy on tact tests following listener training. For the remaining children, some degree of interdependence between the two repertoires was noted. Overall, results of the three studies indicated a greater efficiency of tact than listener training in terms of generating untrained relations (see Table 3). In Watters et al.'s (1980) study, there was a clear indication for all participants that tact training was more likely than listener training to generate the untrained repertoire. In the other two studies, the results were somewhat inconsistent both across and within participants in that respect. Watters et al. additionally reported that tact training consumed fewer trials than listener training.

Results for other children. Children under the age of six without diagnoses of autism or mental retardation participated in 13 studies. Across those studies, 73 children received listener training followed by tact tests, and 46 children received tact training accompanied by testing of listener relations (see Table 1). Most of the children were reported to be typically developing, except that the children in Connell's (1986) study

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

were diagnosed with language disorders. In addition, the children in Williams and McReynolds's (1975) study had articulation difficulties.

As shown in Table 3, complete functional independence of tact and listener repertoires was demonstrated with two participants (Connell & McReynolds, 1981; Williams & McReynolds, 1975). Fifteen additional children participated in within-subject evaluations of both types of training. Only one child passed both tact tests following listener training and listener tests following tact training (Connell & McReynolds). For 10 children, tact training appeared to generate listener relations more readily than listener training generated tacts (Connell & McReynolds; Eikeseth & Jahr, 2001; Holdgrafer & McReynolds, 1975; Williams & McReynolds), and there was tentative evidence of this as well in the first experiment of the Lipkens et al. (1993) single-case study. The opposite pattern was reported for only one child (Eikeseth & Jahr). Thus, the results from the within-subject evaluations are largely consistent with those reported for participants with developmental disabilities.

Also consistent are the results from the studies by Horne, Lowe, and colleagues (Horne et al., 2004, 2006; Lowe et al., 2002, 2005), in which tact training consistently generated listener relations for all participants, but listener training generated tacts less reliably. This effect was, however, not consistently obtained in other studies in which different participants received tact and listener training. Cuvo and Riva's (1980) participants' performance on untrained relations did not differ significantly between tact and listener training conditions, and was fairly accurate on the average. By contrast, none of Connell's (1986) tact-trained participants acquired listener relations, and none of the listener-trained participants acquired tacts. Similarly, initial training of either tact

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

or listener relations did not generate untrained relations to criterion for a majority of the participants in the Y. Barnes-Holmes et al. (2001a, 2002b) studies. However, after the listener-trained participants received tact training on one or two stimulus sets, and the tact-trained participants received listener training on one or two sets, subsequent tact and listener training on new sets began to reliably generate the untrained repertoire. The authors hypothesized that rather than reflecting functional independence of the two repertoires, the participants' initial performance reflected a lack of history with deriving new relations in the presence of the contextual cues contained in the experiment.

As Table 3 shows, when the overall performance from all studies is aggregated, once again a greater proportion of participants consistently passed all listener tests following tact training (50.0%) than tact tests following listener training (27.4%), and a greater proportion failed tact (34.2%) than listener tests (15.2%). A different pattern emerges when terminal performance is viewed, as listener training overall appears to have a greater effect than tact training on the emergence untrained relations. However, this pattern may largely be attributed to the Y. Barnes-Holmes et al. (2001a, 2001b) studies, in which variables were systematically manipulated in order to bring about the emergence of the untrained repertoire, and the training and testing of each child continued until this effect was achieved. Because in those studies, 24 children received listener training but only 8 received tact training, their results inflate the proportion of participants passing tact tests following listener training.

As for other findings, Holdgrafer and McReynolds (1975) reported that although tact training reversed listener relations and vice versa for all of their participants, the effects of either type of training did not generalize to novel exemplars, whether the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

trained or untrained repertoire was tested. By contrast, Connell (1986) reported tact generalization following tact training, but no listener generalization following listener training. In a related vein, Connell and McReynolds (1981) found that tact training generated derived listener relations for 3 out of 6 participants, while listener training never generated derived tacts. Additionally, Cuvo and Riva (1980) and Connell and McReynolds reported fewer trials to criterion in tact than in listener training; a difference that was statistically significant in both cases.

Conclusion. Many of the studies reviewed reported inconsistent effects of tact and listener training, both within and across participants. That is, participants often passed tests of the untrained repertoire following some instances of training but not others, and each type of training often had differential effects across participants. This indicates that a fruitful avenue of research might be to search for variables that affect the emergence of one repertoire as a result of the training of the other. Pre-existing verbal repertoires are likely to be an important variable, as evidenced by the fact that overall a greater degree of functional interdependence was observed among typically developing children. However, the existing verbal repertoires of the participants in the studies reviewed here are difficult to evaluate, as results of language or intellectual assessments were inconsistently reported across studies. One variable that did seem to affect the emergence of untrained repertoires, and was systematically manipulated in two studies (Y. Barnes-Holmes et al., 2001a, 2001b) was repeated exposure to the training and testing procedures employed in the study. Specifically, when participants had been exposed to both listener and tact training contingencies, additional listener or tact training in some cases sufficed to established an untrained tact or listener repertoire

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for some or all participants (Y. Barnes-Holmes et al., 2001a, 2001b; Guess & Baer, 1973; Horne, 2004; Lee, 1981; Smeets, 1978; Smeets & Striefel, 1976).

Although the results of many of the studies were inconsistent within or across participants, a fairly consistent pattern emerges when the results are viewed across studies, in spite of the heterogeneity of participant characteristics and procedures employed. While functional independence of tact and listener repertoire was occasionally observed, in most cases some degree of interdependence appeared to exist. Overall, tact training appeared to generate listener relations more consistently than listener training generated tacts, and an additional benefit of tact training sometimes reported was a saving of training trials relative to listener training.

From the point of view of Skinner's (1957) analysis, it is not clear why tacts should more readily generate listener relations than vice versa. Neither would such results necessarily be expected from the point of view that tacts and listener relations are mutually entailed (D. Barnes-Holmes et al., 2000). Horne and Lowe (1996; p. 202; also see Horne et al., 2004), however, have suggested that young children may acquire what they term name relations, or bi-directional tact and listener relations, more readily as a result of tact reinforcement than the reinforcement of listener behavior. In their view, this may be the case due to beginning language learners' limited echoic repertoires. In order for listener training to generate a tact, Horne and Lowe's analysis assumes that during training, the learner needs to emit an overt or covert echoic response to the verbal stimulus in the presence of the relevant nonverbal stimulus. For tact training to generate a listener relation, by contrast, they suggest that the learner may need only to look at the tacted stimulus in the presence of the vocal (or other) stimulus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

generated by his or her own tact, at the time of reinforcement. Thus, no echoic repertoire is required. Horne and Lowe's analysis assumes that once an extensive echoic repertoire is established, children acquire generalized naming skills, as a result of which the reinforcement of either a tact or a listener relation will almost inevitably establish both relations. However, as Michael (1996) has pointed out, it is also possible to consider the component repertoires of naming relations separately without assuming naming as a higher-order operant as do Horne and Lowe. From that perspective, the extent to which tact training generates listener relations, or vice versa, always depends on the extent to which the relevant collateral responses in fact do occur prior to reinforcement of an overt response. It might be hypothesized that some such collateral responses occur more readily than others. For example, a child may be more likely to look at the tacted stimulus during tact training than he is to make an echoic response to a verbal stimulus during listener training. If that is the case, tact training might be more likely to establish listener relations than listener training to establish tacts, even in the presence of an accurate echoic repertoire.

In any case, the results of the studies that have investigated the effects of both tact and listener training appear to have certain implications for sequencing of programs in EIBI curricula. Although the effects were not entirely consistent across studies and participants, the results appear to suggest that more children would benefit from initial tact than listener training. Additionally, they may suggest that if a child does not readily acquire one repertoire as a result of the training of the other, it may be beneficial to expose the child to both types of training at the same time, following which the training of one relation may be more likely to establish both. In other words, it appears that the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

recommendation to teach receptive before expressive skills (Leaf & McEachin, 1999; Maurice et al., 1996) is not necessarily warranted in the case of tact training as an instance of expressive training. However, many expressive programs target other verbal operants, such as the intraverbal.

Listener Behavior and the Intraverbal

A number of programs in EIBI curricula include both listener and intraverbal training components; for example, programs that teach children to categorize items in terms of their functions, features, or class. As with tact and listener training, completion of listener categorization programs is often considered a prerequisite to teaching intraverbals or other expressive categorization skills (Leaf & McEachin, 1999; Maurice et al., 1996).

Due to the absence of a nonverbal stimulus in a pure intraverbal relation, it intuitively appears unlikely that appropriate listener behavior would emerge as a result of intraverbal training alone, or vice versa. For example, a child who has learned to say "cat" in response to the instruction /Name an animal./ would probably be unlikely to select a picture of a cat when told either /Show me the cat./ or /Show me the animal./, in the absence of any additional training involving cats as visual stimuli. However, if the child were previously able to either point to a cat given its spoken name, or to tact it as "cat", those relations might serve to link a second listener relation involving a cat as a nonverbal stimulus with an intraverbal relation involving either the verbal stimulus /cat/, or "cat" as a response topography.

In contrast to the relatively large body of literature that exists on the effects of tact and listener training, it appears that no published studies have evaluated the effects

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of both intraverbal and listener training on the emergence of untrained listener and intraverbal repertoires. Only one study (Luciano, 1986) appears to have directly addressed the question of whether the training of an intraverbal repertoire may result in new listener relations. Luciano taught three adolescents diagnosed with mental retardation to respond intraverbally to questions about members of categories, such as "What are some foods?". The participants were already able to tact a variety of exemplars from each category. In baseline, all participants were also able to respond as listeners to instructions to select exemplars from some of the categories (e.g., /Give me foods./); however, correct intraverbal responses to the spoken category names were rarely observed. For other categories, this type of listener behavior was not present in baseline, but emerged as a result of intraverbal training. Intraverbal training thus established some new listener relations, whereas the presence of relevant listener relations in baseline was not accompanied by the appropriate intraverbal repertoire.

A few other studies have indirectly produced findings consistent with those obtained by Luciano (1986). First, C. T. Sundberg and Sundberg (1990) conducted a study designed to compare selection-based and topography-based verbal behavior in terms of acquisition speed and the formation of equivalence classes. The participants were four adults with mild or moderate mental retardation. Two relations were trained: tacts of novel objects and topographically identical intraverbal responses to spoken nonsense words. In the topography-based condition, the target tacts and intraverbal responses consisted of nonsense signs. Following training, three of the participants were tested for the emergence of listener relations; that is, the ability to select the appropriate objects given the spoken words. Correct selections emerged for all three participants in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

this condition, although one of them performed below criterion level. The training of topographically identical tact and intraverbal relations thus appeared to generate new listener relations. A similar finding was obtained by Remington and Clarke (1993b). This study extended the authors' prior research (Clarke, Remington, & Light; 1986; Remington & Clarke, 1993a) on *total communication training*; a procedure in which a speaker is trained to emit a sign when simultaneously presented with a nonverbal stimulus and a spoken word, similar to the expressive simultaneous communication training procedure evaluated by Watters et al. (1981). Results of those prior studies indicated that the extent to which total communication training produced appropriate selection of nonverbal stimuli in response to spoken words depended on the extent to which the spoken word had acquired stimulus control over signing. That is, the emergence of those listener relations depended on whether or not training had successfully established an intraverbal relation. Remington and Clarke (1993b) then evaluated the effects of interspersing total communication trials with trials on which the spoken word was presented without the nonverbal stimulus, thus ensuring that intraverbal relations were established. Their participants were children with developmental disabilities, five of whom received the experimental treatment. Following training, three of the children passed tests of listener responding in the presence of spoken names, and an additional child showed considerable improvement. The only child who showed little or no improvement also performed poorly on stimulus control tests over signing on which only the verbal stimulus was presented, and thus the intraverbal training trials appeared not to have successfully established an intraverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

repertoire. The remaining four children had all acquired both tact and intraverbal repertoires.

Taken together, the results of Luciano (1986), C. T. Sundberg and Sundberg (1990), and Remington and Clarke (1993b) indicate that when an identical response topography has been trained both as a tact and an intraverbal, such that a nonverbal and a verbal S^D are functionally equivalent, untrained listener behavior frequently emerges with respect to the verbal stimulus. No research, however, appears to have evaluated whether training the opposite intraverbal relation, one that does not share a common response topography with the relevant tact (e.g., saying "animal" in response to /What is a cat?/, being previously able to tact a cat as "cat") would also result in untrained listener behavior (e.g., responding correctly to /Show me an animal./).

As for the effects of listener training on intraverbal responding, as noted earlier, Luciano (1986) showed that baseline presence of relevant listener behavior was not accompanied by intraverbals. Apart from that evidence, little is known about the extent to which training listener behavior may result in the emergence of intraverbal responding. Only one study (Miguel, Petursdottir, & Carr, 2005) appears to have directly evaluated the effects of listener training on intraverbal responding. However, additional evidence may perhaps be obtained from two studies that evaluated the effects of tact training on intraverbals (Lipkens et al., 1993; Partington & Bailey, 1993), assuming that tact training successfully established some listener relations as well. First, in Experiment 3 of the previously described Lipkens et al. (1993) study, the participant was taught to emit two different tacts when presented with the same picture and asked either /What is this?/ or /What does this say?/. Intraverbals were then tested by asking

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

/What does [name] say?/ or /[sound] what do you hear?/. Correct intraverbals were not observed initially. However, intraverbals gradually emerged over the course of testing, during which test trials were interspersed with similar trials using familiar animal names and sounds. Second, Partington and Bailey (1993) evaluated the effects of tact training on the acquisition of intraverbals in typically developing preschoolers. In Experiment 2, their participants learned to tact items displayed on pictures by saying both the names of the items and the names of the categories to which they belonged (e.g., "It is an apple and it is a fruit."). Intraverbal responding was then tested by asking the children to name category exemplars (e.g., /What are some fruits?/). Although some intraverbal responding emerged for 2 of the 4 children, no child responded intraverbally with all of the topographies that had been trained as tacts until intraverbal responding was directly trained.

Miguel et al. (2005) replicated Partington and Bailey's (1993) results with similar participants, and additionally assessed the effects of listener training, in which the participants were taught to select pictures of items when presented vocally either with their names or their category names. For three children, listener training was conducted on two categories at the beginning of the study. The effects on intraverbal responding differed across children; one child's performance on intraverbal probes improved in both categories, another child showed consistent improvement in only one category, and the third child in neither category. In no case, however, did performance on intraverbal probes meet criterion level until the intraverbal repertoire was directly trained. Three other children initially received multiple-tact training on two categories, followed by listener training. Although two children's performance on intraverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

probes had previously increased as a result of multiple-tact training, none of the three showed additional improvement following listener training.

So far, the limited evidence appears to indicate that as with tacts and listener behavior, listener training may have less of an effect on intraverbal responding than intraverbal training on listener behavior. However, no research appears to have systematically evaluated the effects of both types of training.

The Present Study

The purpose of the present study was to provide an experimental evaluation of the effects of listener training and intraverbal training on the acquisition of listener and intraverbal repertoires. The evaluation was conducted in the context of teaching typically developing children to categorize arbitrary stimuli. Initially, the children were taught to tact a number of category exemplars, as well as to respond as listeners by selecting the nonverbal stimuli in response to their spoken exemplar names. Following baseline measures of listener and intraverbal categorization, each child was exposed to categorization training. Three children received listener categorization training, in which they were trained to respond to category names by selecting appropriate category exemplars. Three others received intraverbal training, in which they were taught to say appropriate category names in response to exemplar names. The target intraverbal repertoire, therefore, was one in which the S^D, rather than the response, corresponded to a previously trained tact (as when a child who has learned to tact a cat responds to the question /What is a cat?/ by saying “An animal.”). Pre- and post-tests were used to assess collateral effects on other categorization skills commonly included in EIBI curricula. Those skills included (a) category matching; that is, selecting a nonverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stimulus when presented with another nonverbal stimulus from the same category (as when a child selects a picture of a dog given a picture of a cat); (b) category tacts; that is, saying a category name rather than an exemplar name when presented with a nonverbal stimulus (as when a child says “animal” rather than “cat” upon seeing a cat), and (c) reverse intraverbals; saying exemplar names when presented with a category name (as when a child says “Cat, dog” when asked to /Name some animals./).

METHOD

Participants and Setting

Six children participated in the study; two girls and four boys. As shown in Table 4, the children’s ages ranged from 3 yrs., 2 mos. to 3 yrs., 11 mos. at the beginning of the study, and from 3 yrs., 5 mos. to 4 yrs. 2 mos. at its completion. All of the children spoke English as a native language, and none of them were reported by their parents to have any known developmental delays. The children were recruited from a local child-care center that gave permission for the study to be conducted within its facility. Prior to each child’s entry into the study, parent permission and child assent were obtained. Participant recruitment and experimental procedures were approved by Western Michigan University’s Human Subjects Institution Review Board (see Appendix A).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4
Participants' Ages in Years (Months) at Entry and at the Beginning of Each Testing Condition.

	Entry	Baseline	Post-testing Set 1	Post-testing Set 2
Marc	3 (5)	3 (7)	3 (8)	3 (8)
Greg	3 (4)	3 (5)	3 (6)	3 (8)
Sam	3 (11)	4 (0)	4 (0)	
Erika	3 (8)	3 (9)	3 (11)	4 (2)
Chanele	3 (2)	3 (3)	3 (4)	
Sanjay	3 (5)	3 (7)	3 (8)	

Experimental sessions were conducted in a partitioned area in the corner of a large room at the child-care center. Other children were frequently present and engaged in regular activities outside of the partitioned area; thus noise level varied from one session to another. During all sessions, the participant and the experimenter were seated side by side at a child-sized table. During some sessions, a second observer was present and seated at the end of the table, on the other side of the child from the experimenter.

For each child, sessions were conducted up to two times a day, five days a week. The weekly frequency of sessions, however, varied depending on the children's attendance at the child-care center and their willingness to attend sessions, as well as on experimenter availability. Each session lasted approximately 10-20 min. During pretraining, exemplar name tact and listener training, the extinction (*EXT*) condition of categorization training, and all testing conditions, each session consisted of a predetermined number of trials blocks, the number and length of which varied depending on the experimental condition. In order to allow the child to monitor his or her progression through the session, the experimenter placed a colored block on a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

similarly colored square that was drawn on a piece of paper, following the completion of approximately 20% of the scheduled trials. The session ended when five blocks had been placed on their corresponding squares. During the continuous reinforcement (*CRF*) condition of categorization training (except for Greg and Chanele's first five sessions on each set), by contrast, the end of a session was marked by the child's filling up a token board. At the end of each session, the child was allowed to select a small toy to take home or a small snack to consume before returning to the classroom. This reward was contingent only on completing the entire session, and not contingent on performance in the session; however, the length of the session could vary with performance when the token board was in use.

Stimuli

Visual stimuli consisted of framed line drawings of twelve outline maps of foreign countries (eight countries in Africa and four additional countries), and twelve characters from foreign writing systems (two of each from Greek, Cyrillic, Hiragana, and Katakana, and four from Hebrew). The drawings were printed in black ink on white 21.6 x 27.9 cm sheets of paper (*trial sheets*), as well as on 5 x 5 cm white cards (*stimulus cards*). All trial sheets were inserted into plastic sheet protectors contained in a three-ring binder, and stimulus cards were inserted into transparent hard plastic covers measuring 7.5 x 10 cm. In most training and testing conditions in which visual stimuli were used, one sheet was presented on each trial. Depending on the specific training or testing condition, each trial sheet contained either one stimulus centered in the middle of the sheet (tact trials), three vertically aligned stimuli (listener trials), or one sample stimulus centered in the upper half of the sheet, and three vertically aligned comparison

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stimuli in the bottom half of the sheet, covered by a paper flap (category matching trials). Following the completion of a trial, the page was turned over, and a blank sheet appeared. When the blank sheet was turned over, the next trial sheet appeared. Stimulus cards were used instead of trial sheets during the continuous reinforcement (*CRF*) conditions of listener categorization training.

The eight maps of countries in Africa were divided into four 2-member categories labeled “north”, “south”, “east”, and “west”. Eight of the foreign characters were similarly divided into four 2-member categories, labeled “Greek”, “Cyrillic”, “Kata”, and “Hira”. Thus a total of eight categories containing a total of 16 exemplars were used in the experiment, but each child received training on a maximum of four categories. The eight categories were divided into four category sets. Each set contained either two map categories or two character categories. Each exemplar was assigned a specific name, which was either the conventional name of the foreign character or the country represented by the map, or a modified form of its conventional name. Figure 1 shows the visual stimuli in each category set along with their exemplar names.

The four additional maps and the four Hebrew characters served as one of two negative comparisons on listener categorization and category matching trials. As shown in Figure 2, those stimuli were also assigned specific names based on their conventional names, but they were not assigned category membership.

In addition to maps and characters, line drawings of two familiar animals (cat and dog), two familiar fruits (apple and banana) and two other familiar items (a shoe and a flower) were used in the pretraining condition of the experiment (see Appendix B).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

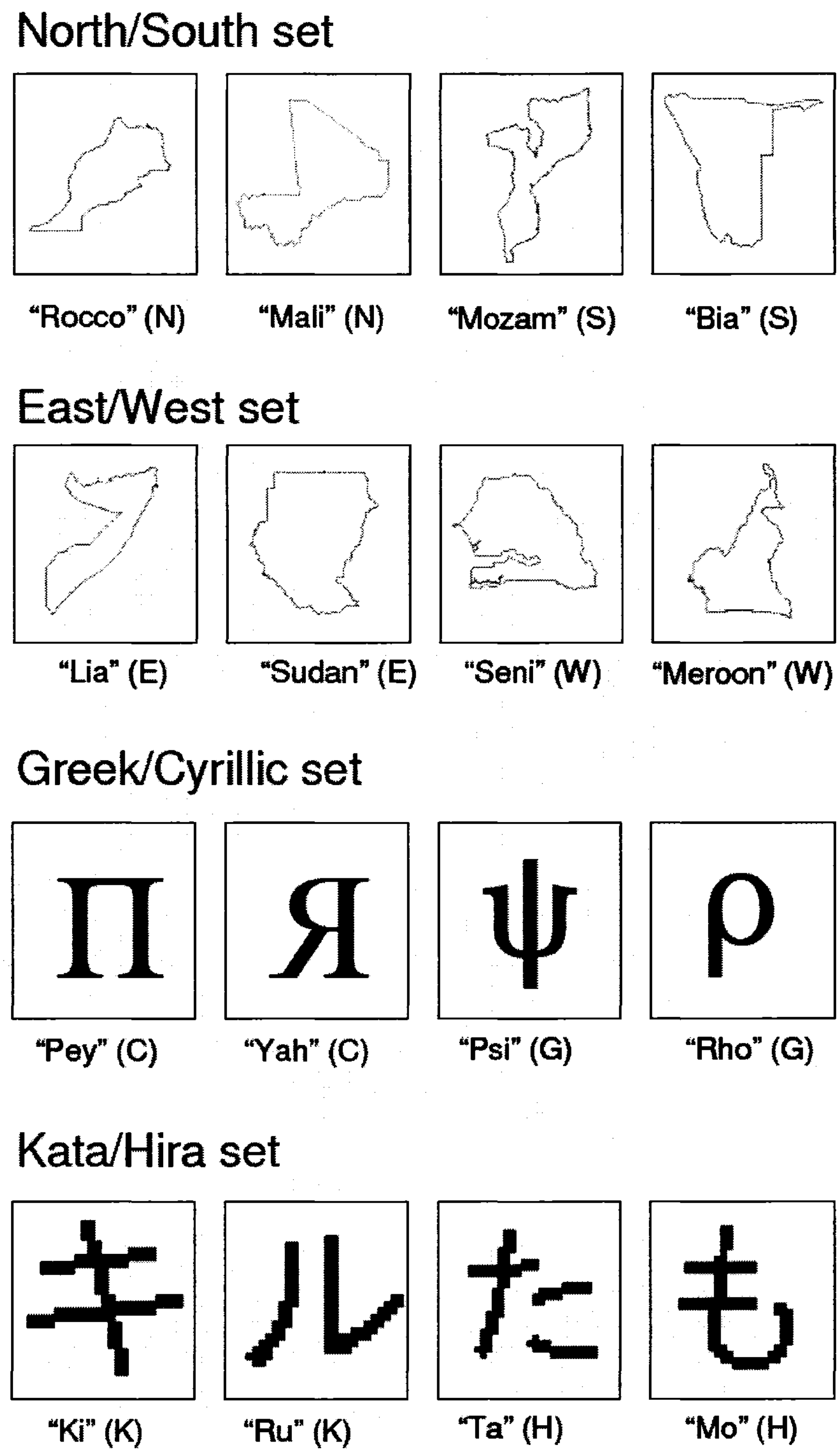


Figure 1. Category sets and exemplar names.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

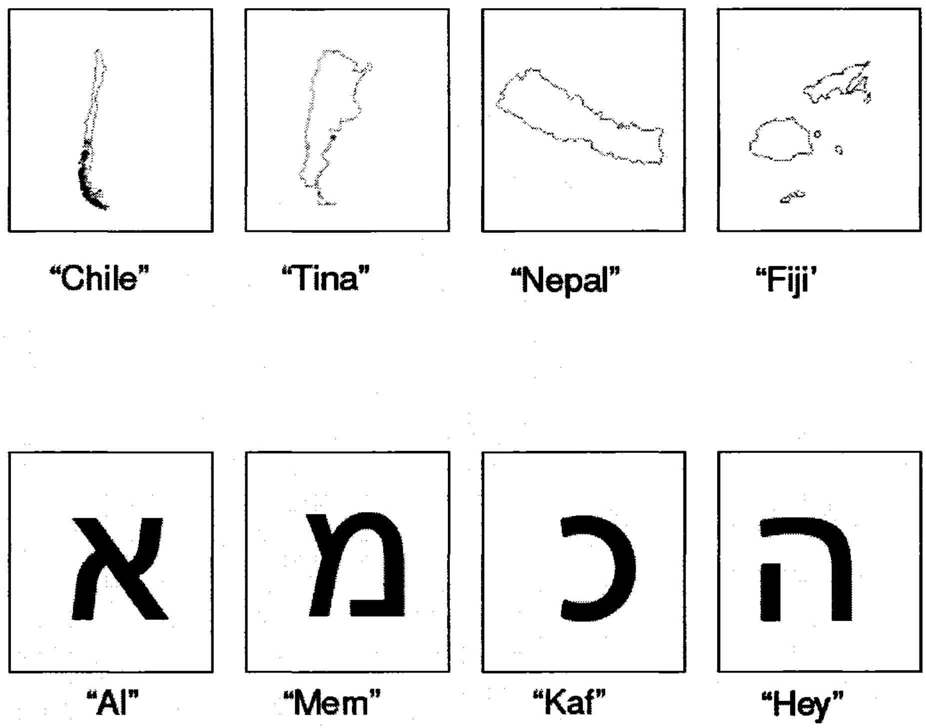


Figure 2. Additional stimuli used as comparisons on listener categorization and category matching trials.

Data Collection

Dependent variables. The primary dependent variables were (a) listener categorization; defined as selecting a visual stimulus given its category name, and (b) intraverbal categorization; defined as saying the name of a category given the specific name of one of its exemplars. Secondary dependent measures were performance on pre- and post-tests of (a) category matching; defined as selecting a visual comparison stimulus that belonged to the same category as a visual sample stimulus, (b) category tacts; defined as saying a category name when presented with a visual stimulus from

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that category; and (c) reverse intraverbal categorization; defined as saying an exemplar name given a category name. In addition, data were collected on collateral overt verbal responses that occurred during intraverbal and listener categorization training and testing.

Response measurement. During all training and testing conditions, the experimenter recorded a correct or an incorrect response on a data sheet for each trial. On listener trials and category matching trials, a response was scored as correct if the child touched the experimentally defined correct stimulus within 10 s of the presentation of comparison stimuli. An incorrect response was recorded if the child touched another stimulus or did not touch a stimulus within 10 s. If the child touched more than one stimulus, only the first response was recorded. On all trials requiring a vocal response, a correct response was recorded if the child vocalized the experimentally defined correct category name or exemplar name. An incorrect response was recorded if the child did not emit the target response within 10 s of the experimenter's instruction, or if the child said the name of a different category (when the target response was a category name), or a different exemplar (when the target response was an exemplar name). If the child said the name of more than one category (when the target response was a category name), or more than one exemplar (when the target response was an exemplar name), only the first response was typically scored as correct or incorrect. On test trials of reverse intraverbal categorization, however, the first two responses were recorded and each scored as correct or incorrect.

Data on collateral verbal responses during listener and intraverbal categorization training and testing were collected as follows. During all listener categorization trials,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

under both training and testing conditions, an echoic response was recorded on the data sheet if the child repeated the category name contained in the experimenter's instruction at any time before the stimuli were removed, and a tact was recorded if the child vocalized the exemplar name of the stimulus that he or she pointed to. During all intraverbal categorization trials, under both training and testing conditions, an echoic response was recorded if the child repeated the exemplar name contained in the experimenter's instruction before responding with the category name, and a reverse intraverbal was recorded if the child vocalized the exemplar name after responding with a category name.

Interobserver agreement. A second observer independently collected data on at least 25% of all trial blocks or tests for each child in each categorization training and testing condition. For target responses, an agreement was scored for each trial in which the experimenter and the observer both recorded a correct or an incorrect response; otherwise a disagreement was scored. Point-by-point overall agreement was calculated for each trial block or test by dividing the number of agreements by the sum of agreements and disagreements and multiplying by 100%. Table 5 shows the percentage of trial blocks or tests for which agreement was assessed, along with average agreement on target responses across trial blocks or tests, during categorization training, listener and intraverbal categorization testing, and pre- and post-tests. In each condition, average agreement exceeded 97%. Table 5 also shows the range of agreement scores across trial blocks. During categorization training, each trial block consisted of only 4 trials, and thus 50% agreement was recorded for blocks on which the two observers disagreed on two responses, and 75% agreement if they disagreed on one response.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5
Percentage of Trial Blocks or Tests on Which Interobserver Agreement Was Assessed, Mean Agreement, and Range of Agreement for Each Participant During Categorization Training and Testing.

	Categorization Training		Listener and intraverbal categorization testing		Pre- and post-tests	
	Percent (%) blocks assessed	Mean % agreement (range)	Percent (%) blocks assessed	Mean % agreement (range)	Percent (%) blocks assessed	Mean % agreement (range)
Marc	69.4	100.0	86.0	100.0	58.3	100.0
Greg	60.9	99.5 (50.0-100.0)	57.6	99.3 (87.5-100.0)	60.0	98.9 (90.0-100.0)
Sam	27.6	99.1 (75.0-100.0)	55.6	97.5 (87.5-100.0)	50.0	98.3 (95.0-100.0)
Erika	49.6	99.8 (75.0-100.0)	57.7	99.2 (87.5-100.0)	83.3	100.0
Chanele	56.2	98.9 (75.0-100)	60.0	100.0	33.3	100.0
Sanjay	92.5	100.0	100.0	100.0	83.3	99.0 (95.0-100.0)

For collateral verbal responses during categorization training, point-by-point occurrence agreement rather than overall agreement was calculated. For most of the children, such collateral responses were rarely observed, and thus including agreement on nonoccurrences would have inflated the agreement percentage. An agreement was scored when the experimenter and the second observer both recorded a specific collateral response, and a disagreement was scored if only one of them recorded a response. For Greg, collateral responses were observed on 10 trials. Agreement data were available on 6 of those trials (60%), and occurrence agreement was 83.3%. Corresponding data for Sam were 162 trials, 38 (23.5% of trials), and 81.6%; for Erika, 5 trials, 1 (20.0% of trials), and 0%; and for Chanele, 16 trials, 4 (25.0% of trials), and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

75.0%. For Marc, only one collateral response was recorded and only by one of the two observers present. For Sanjay, no collateral responses were recorded.

Procedural fidelity. For each child, procedural fidelity was assessed on at least 19% of all categorization training blocks, and at least 20% of all blocks of intraverbal and listener categorization testing. To assess procedural fidelity, an independent observer recorded antecedents (correct or incorrect instruction) and consequences (praise, error correction, or no consequence) delivered by the experimenter on each trial. A trial was scored as correctly implemented if the experimenter delivered the instruction correctly and delivered consequences as dictated by the child's response and the experimental condition. A procedural fidelity score was computed for each block by dividing the number of correctly implemented trials by the total number of trials, and multiplying by 100%. Almost all scored trials were correctly implemented. Table 6 shows the percentage of blocks on which procedural fidelity was assessed for each child in each condition, as well as average fidelity scores.

Table 6
Percentage of Trial Blocks or Tests on Which Procedural Fidelity Was Assessed and Average Fidelity During Categorization Training and Testing.

	Categorization Training		Listener and Intraverbal Categorization Testing	
	Percent (%) blocks assessed	Mean % fidelity score	Percent (%) blocks assessed	Mean % fidelity score
Marc	24.0	100.0	39.5	100.0
Greg	29.3	99.6	30.3	100.0
Sam	19.1	100.0	55.6	100.0
Erika	20.3	100.0	30.7	100.0
Chanele	24.0	100.0	30.8	100.0
Sanjay	54.7	100.0	20.0	100.0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Procedure

Experimental design. Each child received either intraverbal (Marc, Greg, Sam) or listener (Erika, Chanele, Sanjay) categorization training on either one or two category sets, preceded and followed by baseline and post-training measures of listener and intraverbal categorization. The effects of each type of training on those primary dependent variables were evaluated either in a multiple-baseline design across category sets (Marc, Greg, Erika) or in a nonconcurrent multiple-baseline design across participants (Chanele and Sanjay). Sam's participation in the experiment was terminated before it was possible to complete training and testing procedures on the first category set, but his baseline and initial post-training data are presented as an A-B design. For each category set, pre-test and post-test performance was used to evaluate the effects of training on category matching, category tacts, and reverse intraverbal categorization.

Table 7 presents an overview of all experimental conditions in the order that they were conducted for participants who received training on two category sets. Conditions were identical for participants who received training on only one set, except that no training or post-test procedures were implemented for the second set, and for Sanjay, no baseline measures were conducted on the second set.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 7

An Overview of Experimental Conditions for Participants Who Received Categorization Training on Two Category Sets.

Phase	Condition	Criterion for Completion
Preexp. echoic assessm.		
Pretraining	Pretraining: Tact Training – Exemplar Names	1 x 4 trials at 100%
	Pretraining: Listener Training – Exemplar Names	1 x 4 trials at 100%
	Pretraining: Listener Categorization	1 x 4 trials at 100%
	Pretraining: Intraverbal Categorization	1 x 4 trials at 100%
	Pretraining: Category Matching	1 x 4 trials at 100%
	Pretraining: Category Tacts	1 x 4 trials at 100%
	Pretraining: Reverse Intraverbal Categorization	1 x 4 trials at 100%
Tact/Listener Training	Tact Training – Exemplar Names: Group 1	1 x 12 trials at 100%
	Tact Training – Exemplar Names: Group 2	1 x 12 trials at 100%
	Tact Training – Exemplar Names: Group 3	1 x 12 trials at 100%
	Tact Training – Exemplar Names: All stimuli	3 x 12 trials at 100%
	Listener Training – Exemplar Names	First 12 trials at 100% or
		3 x 12 trials at 100%
*Pre-testing	Pretraining Review: Category Matching	1 x 4 trials at 100% EXT
	Category Matching Pre-test: Set 1	20 trials EXT
	Category Matching Pre-test: Set 2	20 trials EXT
	Pretraining Review: Category Tacts	1 x 4 trials at 100% EXT
	Category Tact Pre-test: Set 1	20 trials EXT
	Category Tact Pre-test: Set 2	20 trials EXT
	Pretraining Review: Reverse Intraverbals	1 x 4 trials at 100% EXT
	Reverse Intraverbal Pre-test: Set 1	20 trials EXT
	Reverse Intraverbal Pre-test: Set 2	20 trials EXT
*Baseline	Pretraining Review: Listener Categorization	1 x 4 trials at 100% EXT
	Pretraining Review: Intraverbal Categorization	1 x 4 trials at 100% EXT
	Listener/Intraverbal Categorization Baseline: Sets 1 & 2	Stable zero or chance level responding EXT
*Categorization Training: Set 1	*CRF Stage 1* (Erika, Sanjay, Sam, Marc)	3 x 4 trials at 100%
	CRF Stage 2 (Erica, Sanjay, Sam, Marc)	3 x 4 trials at 100%
	CRF Stage 3 (All participants)	5 x 4 trials at 100%
	Extinction	3 x 4 trials at 100% EXT

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3–Continued

Phase	Condition	Criterion for Completion
	Listener/Intraverbal Categorization Baseline: Set 2 (concurrent with training on Set 1)	
*Post-testing: Set 1	Pretraining Review: Listener Categorization	1 x 4 trials at 100% EXT
	Pretraining Review: Intraverbal Categorization	1 x 4 trials at 100% EXT
	Listener/Intraverbal Categorization Post-testing: Set 1	Success or failure criterion EXT
	Pretraining Review: Category Matching	1 x 4 trials at 100% EXT
	Category Matching Post-test: Set 1	20 trials EXT
	Pretraining Review: Category Tacts	1 x 4 trials at 100% EXT
	Category Tact Post-test: Set 1	20 trials EXT
	Pretraining Review: Reverse Intraverbal	1 x 4 trials at 100% EXT
	Reverse Intraverbal Post-test: Set 1	20 trials EXT
	Listener/intraverbal Categorization Baseline: Set 2	Stable zero or chance level responding EXT
*Categorization Training: Set 2	*CRF Stage 1* (Erika, Marc)	3 x 4 trials at 100%
	CRF Stage 2 (Erika, Marc)	3 x 4 trials at 100%
	CRF Stage 3 (Erika, Marc, Greg)	5 x 4 trials at 100%
	Extinction	3 x 4 trials at 100% EXT
*Post-testing	Pretraining Review: Listener Categorization	1 x 4 trials at 100% EXT
	Pretraining Review: Intraverbal Categorization	1 x 4 trials at 100% EXT
	*Listener/intraverbal Categorization Post-tests: Set 2	Success or failure criterion EXT
	Pretraining Review: Category Matching	1 x 4 trials at 100% EXT
	Category Matching Post-test: Set 2	20 trials EXT
	Pretraining Review: Category Tacts	1 x 4 trials at 100% EXT
	Category Tact Post-test: Set 2	20 trials EXT
	Pretraining Review: Reverse Intraverbal	1 x 4 trials at 100% EXT
	Reverse Intraverbal Post-test: Set 2	20 trials EXT

Note. The implementation of the independent variable is italicized. An asterisk before a condition name indicates that maintenance trials are conducted throughout the phase for specific name tacts and listener behavior. EXT = under extinction.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Pre-experimental echoic assessment. The first time a participant met with the experimenter, the experimenter presented an echoic trial for each response topography trained and tested in the experiment. On an echoic trial, the experimenter said, for example, "Can you say Mali?" and waited for a response. Correct echoic responses were praised. If the child initially did not make an accurate echoic response, the trial was repeated twice. Failure to respond accurately on each of the two additional trials would have resulted in the response topography being replaced with one to which the child could respond. However, this was unnecessary.

Pretraining with familiar categories. Before the experiment began, pretraining with two exemplars from each of two familiar categories (animals and fruits) was conducted to familiarize the children with the experimental procedures, and to ensure instructional control. As shown in Table 7, each child received tact and listener training on exemplar names, listener categorization training, intraverbal categorization training, and training on pre- and post-testing procedures (category matching, category tacts, and reverse intraverbal categorization). In all pretraining conditions, training trials were implemented in four-trial blocks, except that reverse intraverbal categorization pretraining was conducted in two-trial blocks, in which up to two responses were scored as correct on each trial. In each pretraining condition, stimulus presentation and instructions were identical to that in the corresponding experimental conditions described in the following sections.

During all pretraining conditions, all correct responses were praised. If an incorrect response was made or no response was made within 10 s, the experimenter prompted a correct response, verbally or by pointing, and then presented the next trial.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Each pretraining condition continued until the child responded with 100% accuracy in a single trial block.

Prior to each testing condition, the testing procedures used in that condition were reviewed briefly with the familiar categories (see Table 7). In each review session, initially one four-trial block was conducted under extinction. If performance was less than 100% accurate, pretraining was conducted until 100% accuracy was achieved in one trial block, followed by 100% accuracy in one trial block under extinction.

Exemplar name tact training. In this condition, the child was trained to tact each of 12 stimuli; 6 maps and 6 foreign characters. Eight of the stimuli belonged to two different category sets; either Kata/Hira and North/South or Greek/Cyrillic and East/West. The remaining four stimuli consisted of two additional maps and two Hebrew characters, and did not belong to an experimentally defined category. The 12 stimuli were arbitrarily divided into three groups of four stimuli, with each group containing at least one map and at least one character. Training trials were conducted in 12-trial blocks, in which presentation order varied across blocks. Training proceeded through the following stages. In Stage 1, presentations of the four stimuli in one stimulus group were alternated until responding was 100% accurate on one 12-trial block containing three presentations of each stimulus. This procedure was repeated in Stages 2 and 3 for the second and third stimulus groups. Finally, in Stage 4, presentations of all 12 stimuli were alternated until responding was 100% accurate in three consecutive 12-trial blocks, each block containing one presentation of each stimulus.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Training procedures were identical during all stages of training. On each trial, the experimenter asked, “What is this?”, followed by the presentation of a trial sheet containing one stimulus. Correct responses were praised by the experimenter, and the stimulus then removed by turning the trial sheet. When an incorrect response was made, the experimenter prompted a correct response by saying, for example, “This is *Mali*, can you say it?”, and then removed the stimulus.

Exemplar name listener training. Following the completion of exemplar name tact training, each child was trained to respond as a listener to the spoken exemplar names of each of the 12 stimuli that he or she had previously learned to tact. Trials were conducted in 12-trial blocks, with each block containing one presentation of each specific name. On each trial, the experimenter presented an instruction containing the exemplar name (e.g., “Which one is *Mali*?”), and then presented a trial sheet containing 3 of the 12 stimuli; one positive and two negative comparisons. The location of positive comparison stimuli varied across trials, and on each trial, the negative comparison stimuli consisted of one map and one character. Correct responses were praised, following which the experimenter removed the stimuli by turning the trial sheet. If an incorrect response was made, the experimenter pointed to the correct stimulus and said “It is this one”, prompted the child to touch the positive comparison, and then removed the stimuli.

If performance was less than 100% accurate on the first trial block, listener training on exemplar names continued until performance was 100% accurate on three consecutive trial blocks. If performance was 100% accurate on the initial trial block, further listener training was omitted, and pre-testing began.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Exemplar name maintenance. Tacts and listener relations with respect to exemplar names were maintained during all conditions following exemplar name listener training (see Table 7). Maintenance trials for either tacts or listener behavior were conducted immediately prior to approximately every other testing or categorization training session. Maintenance trials were conducted in 12-trial blocks containing one trial per each of the 12 stimuli used in the previous tact and listener training conditions. Approximately half of all maintenance blocks were tact training blocks, and the other half were listener training blocks. Training procedures were identical to exemplar name tact and listener training. If on the first trial block, the child made an error on one or more of the eight trials that targeted category exemplars, additional blocks were conducted until the child responded correctly on all eight trials. Errors made on one of the remaining four stimuli were corrected, but were not followed by additional trial blocks as long as the child responded correctly on all trials with category exemplars.

Category matching pre- and post-tests. A category matching pre-test for each of the two category sets was conducted prior to the listener and intraverbal categorization baseline in each phase. A post-test was then conducted for each set following the post-training listener and intraverbal categorization testing condition for that set. Each test consisted of 20 randomly ordered trials, in which each stimulus in the category set served as a sample five times. The experimenter initiated each trial by presenting a trial page with only the sample stimulus visible to the child, delivering the instruction "Touch this", and waiting for the child to touch the sample stimulus. The experimenter then instructed the child to "Find the one that goes with it." and lifted the flap that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

covered the three comparison stimuli. The correct comparison was a stimulus from the same category as the sample stimulus. The negative comparisons were (a) a stimulus from the other category in the set (e.g., from the "East" category if the sample stimulus was from the "West" category), and (b) a stimulus that the child had been trained to tact but did not belong to an experimentally defined category (a map when map sets were tested, and a Hebrew character when character sets were tested). The experimenter waited up to 10 s for a response and then removed the stimuli by turning the trial sheet. No consequences were provided for correct or incorrect responses.

Category tact pre- and post-tests. As with category matching, category tact pre- and post-tests were conducted prior to baseline and following the post-training testing phase for each set. Each test consisted of 20 randomly ordered trials, in which each stimulus in the category set was presented five times. The experimenter initiated each trial by asking "What is this?", and presenting a trial page containing one stimulus. The child was given up to 10 s to respond. If the child responded with an exemplar name rather than a category name, the experimenter asked "Is it called anything else?", and again waited up to 10 s for a response. No consequences were provided for correct or incorrect responses following either type of instruction. A trial was scored as correct if the child responded with the correct category name following either the first or the second question.

Reverse intraverbal categorization pre- and post-tests. As with category matching and category tacts, pre- and post-tests for reverse intraverbal categorization were conducted prior to baseline and following the post-training testing phase for each set. Each test consisted of 10 randomly ordered trials, five for each category. In each

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

trial, the child could make 0, 1, or 2 correct responses, and thus a total of 20 correct responses were possible. At the beginning of each trial, the experimenter delivered the instruction “Name something that is [category name].”, and waited up to 10 s for a response. If the child responded with two or more exemplar names, or did not respond with an exemplar name, the trial was over and the next trial was initiated. If the child initially responded with only one exemplar name, the experimenter instructed the child to “Name something else that is [category name].” (regardless of whether the first response was correct or incorrect), and again waited up to 10 s for another response before terminating the trial. No consequences were delivered for either correct or incorrect responding following either type of instruction. It should be noted that the second instruction was not presented if, following the first instruction, the child emitted one correct and one incorrect response.

Listener and intraverbal categorization tests. Listener and intraverbal categorization were tested in baseline and immediately following each categorization training phase. One category set was tested at a time. Each trial block consisted of one listener categorization trial and one intraverbal categorization trial for each stimulus in the set, for a total of eight trials. Listener and intraverbal trials were interspersed within trial blocks, and presentation order varied across blocks.

At the beginning of each listener categorization trial, the experimenter presented the instruction “Which one is [category name]?”, and then presented a trial sheet containing one positive and two negative comparison stimuli. As on category matching test trials, one of the negative comparisons was a stimulus from the other category in the set (e.g., from the “East” category, if the correct comparison belonged to the “West”

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

category). The other was a stimulus that the child had been trained to tact but did not belong to an experimentally defined category (a map when map sets were tested, and a Hebrew character when character sets were tested). Following the presentation of the comparison stimuli, the experimenter waited up to 10 s for a response, and then removed the stimuli by turning the trial sheet. No consequences were provided for either correct or incorrect responses.

No visual stimuli were used in intraverbal categorization trials. At the beginning of each intraverbal trial, the experimenter said "[Specific name] is…" (e.g., "*Mali* is…"), and waited up to 10 s for a response. No consequences were delivered for either correct or incorrect responding.

In baseline, testing continued until stable performance was observed on both listener and intraverbal trials. Following training, testing continued at least until performance in the untrained repertoire (listener or intraverbal categorization) met either failure or success criterion, or if neither criterion was met, until responding was stable as determined by visual inspection. Untrained listener and intraverbal categorization were considered successfully acquired if within six blocks of testing, performance was 100% correct (four correct responses) in three consecutive blocks. The failing criterion for intraverbal categorization was three consecutive blocks of 0% correct responding, and the failing criterion for listener categorization was three consecutive blocks of two or fewer correct responses (chance level was 1.33 correct responses per block, as three comparisons were available on each of the four listener trials). Intermediate performances were considered to be evidence of partial acquisition of the untrained repertoire.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

No failure or acquisition criteria were applied to the directly trained repertoire during testing. However, if following training, performance in the trained repertoire showed a consistent decline from training during post-testing as determined by visual inspection, categorization training was reinstated. Following the completion of training, all post-testing procedures were repeated.

Intraverbal categorization training. During intraverbal categorization training, the child was taught to respond to the spoken names of four category exemplars belonging to two different categories, by saying the appropriate category name. The training phase was divided into a continuous reinforcement (*CRF*) condition and an extinction (*EXT*) condition. In both conditions, trials were presented in four-trial blocks.

No visual stimuli were used during intraverbal categorization training. At the beginning of each intraverbal trial, the experimenter said "[exemplar name] is…", and waited up to 10 s for a response. During the CRF condition, all correct responses were praised. In addition, if the child responded correctly on both of a pair of consecutive trials, the experimenter placed a token (coin) on a token board that consisted of a blue piece of paper with 16 squares printed on it. The token board could be folded in such a way that 4, 8, 12, or 16 squares were visible. The token board was used because pilot testing indicated that praise alone might be insufficient to establish stimulus control during intraverbal training. Further, pilot testing indicated that consequating every correct response with a token could establish a pattern of responding in which the child arbitrarily emitted one of the two category names on each trial. This resulted in every other response being reinforced on the average, which appeared to be sufficient to maintain chance responding. Therefore, more than one consecutive correct response

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was required to earn a token. Specifically, trials were presented in pairs, and the child informed that he would earn a token if he answered the next two questions correctly. If one or both responses were incorrect, no token was delivered for that pair of trials, and the child was required to respond correctly on both trials in the next trial pair in order to earn a token. Trial pairs were arranged such that some pairs contained two exemplar names from the same category and the same response was correct on both trials, whereas other pairs contained one exemplar name from each category.

For Greg, the token board and trial-pair procedures was introduced after the first 100 trials on each set, as his performance after 100 trials indicated that he was responding arbitrarily with the two category names at that point. For Marc and Sam, those procedures were introduced immediately at the beginning of training on each set. In order to complete a training session and receive an opportunity to select a reward, the child had to earn 4, 8, 12 or 16 tokens, depending on how many squares were visible on the token board. The number of tokens to be earned in each session was determined by the experimenter prior to the session, based on an informal evaluation of the child's progress in training. Early in training, when two consecutive correct responses were infrequent, only four tokens were required to complete the session. As training progressed, the token board was gradually enlarged, until the child was responding correctly on the majority of all trials, at which time the child was required to earn up to 16 tokens per session. However, fewer tokens were required if session data indicated that the child's performance typically declined from the beginning to the end of a session.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

During the CRF condition, if an incorrect response was made or no response was made within 10 s, the experimenter vocally prompted a correct response (e.g., "East"). If one or more incorrect response was made on a trial pair, then following the presentation of both trials, the entire trial pair was repeated until the child responded correctly on both trials.

For Sam and Marc, the CRF condition was divided into three stages. In Stage 1, presentations of two exemplar names, one from each category, were alternated such that each trial block contained two presentations of each stimulus, but the order of presentation varied across trial blocks. When 100% accuracy was observed on three consecutive trial blocks, Stage 2 commenced, in which presentations of the remaining two stimuli were alternated in the same manner until 100% accuracy was observed on three consecutive blocks. In Stage 3, each four-trial block contained one presentation of each of the four stimuli. Stage 3 training continued until 100% accuracy was achieved in five consecutive trial blocks. For Greg, Stages 1 and 2 were omitted, and all four stimuli presented together in four-trial blocks from the beginning.

Following the completion of the CRF condition, the EXT condition was introduced. The purpose of this condition was to promote maintenance of the trained repertoire under conditions similar to testing. During the EXT condition, all four stimuli were presented in four-trial blocks as in Stage 3 of the CRF condition; however, trials were not presented in pairs. No consequences were delivered following correct responses. If an incorrect response was made, or no response was made within 10 s, the experimenter vocally prompted a correct response, and then presented the next trial. Training under extinction continued until responding was 100% correct on three

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

consecutive trial blocks. If performance declined during this condition, as determined by visual inspection, Stage 3 CRF training was reintroduced, followed by the reinstatement of EXT.

Any time it was necessary to reinstate training due to declining performance in the trained repertoire during post-testing, retraining began at EXT, and the same training criteria were applied as before.

Listener categorization training. During listener categorization training, the child was trained respond to two spoken category names by selecting visual stimuli representing two exemplars from each category. Trials were presented in four-trial blocks. At the beginning of each trial, the experimenter presented the instruction "Which one is [category name]?", and presented the positive comparison along with two negative comparison stimuli. As on listener categorization tests, one of the negative comparisons belonged to the other category in the set and one was a stimulus that the child had learned to tact but did not belong to an experimentally defined category. Listener categorization training, like intraverbal categorization training, was divided into a CRF condition and an EXT condition. During the CRF condition, stimuli were presented on stimulus cards, which were either lined up on the table, or held up by the experimenter in front of the child. For Chanele, trial sheets were used initially, but replaced with stimulus cards after 612 trials of training, because she was failing to make progress. Stimulus cards were subsequently used from the beginning of training for Erika and Sanjay, because experience with Chanele indicated that the greater flexibility in stimulus presentation afforded by the stimulus cards better kept her attention during

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

training. During the EXT condition for all children, the stimulus cards were replaced with trial sheets identical to those used during testing.

To maintain consistency with intraverbal training, listener categorization trials were presented in pairs during the CRF condition. All correct responses were praised, and a token was delivered if the child responded correctly on both trials in the pair. For Chanele, the token board procedure was introduced after the first 100 trials of training revealed that the accuracy of responding did not exceed chance levels. For Erika and Sanjay, the token board procedure was introduced immediately at the beginning of training. The size of the token board in each session was determined in the same manner as during intraverbal categorization training.

If an incorrect response was made during the CRF condition, or no response was made within 10 s, the experimenter pointed to the correct comparison, said "It's this one.", and prompted the child to touch the stimulus. The stimulus materials were then removed and the next trial presented. If one or more incorrect response was made on a trial pair, then following the presentation of both trials, the entire trial pair was repeated until the child responded correctly on both trials.

For Erika and Sanjay, the CRF condition was divided into three stages corresponding to the three stages of intraverbal categorization training for Sam and Marc. The criteria for completing each stage were the same as those used in intraverbal categorization training. For Chanele, Stages 1 and 2 were omitted and training began at Stage 3.

The EXT condition began following the completion of Stage 3 of the CRF condition. At the beginning of this condition, stimulus cards were replaced with trial

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sheets. As in intraverbal categorization training, trials in this condition were presented one by one in four-trial blocks. No consequences were delivered following correct responses. If an incorrect response was made, or no response was made within 10 s, the experimenter prompted a correct response by pointing to the positive comparison stimulus, and then presented the next trial. Training under EXT continued until responding was 100% correct on three consecutive trial blocks. Procedures for reinstating CRF due to declining performance during EXT, as well as for reinstating EXT due to declining performance on subsequent listener categorization test trials, would have been the same as those used for intraverbal categorization training. However, it was never necessary to reinstate either condition for either reason.

RESULTS

Pretraining

None of the children required extensive pretraining on experimental procedures with familiar stimuli; however, they all required more than one block of training in at least one pretraining condition (see Table 8). In most cases, no more than two additional blocks were required, except that Greg required a total of six blocks of training to complete pretraining on reverse intraverbal categorization, and Chanele required 26 blocks to complete category matching. Chanele eventually mastered category matching after the instruction was changed from “Find the one that goes with it.” to “Find the one that matches.”. The new instruction was subsequently used on the category matching pre- and post-tests for Chanele.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As Table 8 also shows, additional training was occasionally required when testing procedures were reviewed with familiar stimuli immediately prior to testing. In all cases, if the child made an incorrect response on the first review block, at least two additional blocks were required to meet the review criterion (one block at 100% accuracy under training conditions and one block at 100% accuracy under extinction), but more than three blocks were required only in one case (category matching for Greg, prior to his category matching post-test on the second set).

Table 8
Number of Trials Required to Complete Each Pretraining and Pretraining Review Condition for Each Participant.

	Marc	Greg	Sam	Erika	Chanele	Sanjay
Initial training trials:						
Tacts: Exemplar names	8	4	8	4	4	4
Listener responding: Exemplar names	4	4	4	4	4	4
Listener categorization	8	8	8	4	4	4
Intraverbal categorization	8	4	4	8	8	4
Category matching	8	4	8	8	104	4
Category tacts	8	4	8	12	8	12
Reverse intraverbals	4	24	4	4	4	4
Review trials: Baseline / Post-tr. Set 1 / Post-tr. Set 2:						
Listener categorization	4 / 4 / 4	12 / 4 / 4	12 / 4	4 / 4 / 4	4 / 4	4 / 4
Intraverbal categorization	4 / 4 / 4	4 / 4 / 4	4 / 4	12 / 4 / 4	4 / 12	4 / 4
Category matching	4 / 4 / 4	4 / 4 / 24	12 / 4	4 / 4 / 4	16 / 4	4 / 4
Category tacts	4 / 4 / 4	4 / 4 / 16	4 / 4	4 / 12 / 4	12 / 12	16 / 4
Reverse intraverbals	4 / 4 / 4	4 / 4 / 4	4 / 4	12 / 4 / 4	4 / 4	4 / 4

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Exemplar Name Training and Maintenance

As shown in Table 9, the number of trials required to complete tact training ranged from 516 (Chanele) to 1788 (Marc). Five of the children responded with 100% accuracy on the first block of exemplar name listener training, and therefore did not require additional listener training. With Erika, listener training was completed in 48 trials.

Table 9
Trials to Criterion in Exemplar Name Training, Average Accuracy on Category Exemplar Trials in the First Maintenance Block in Each Session, and Average Accuracy on Entire First Maintenance Block.

	Trials to criterion		Maintenance: Category exemplars		Maintenance: All stimuli	
	Tacts	Listener relations	Tacts % correct	Listener relations % correct	Tacts % correct	Listener relations % correct
Marc	1788	12	93.8	94.6	92.7	95.2
Greg	900	12	98.2	97.3	98.1	96.9
Sam	612	12	87.5	90.6	91.7	93.8
Erika	960	48	95.3	98.3	96.4	98.9
Chanele	516	12	83.3	97.7	79.9	98.5
Sanjay	1236	12	96.9	100.0	89.6	100.0

Table 9 also shows the average accuracy on the first maintenance block in each session in which maintenance trials were conducted. In most cases, the children maintained over 90% accuracy, on the average, on tact and listener trials that targeted the 8 category exemplars. The same was true of the entire maintenance block, even though correct responding was not required on the four trials that targeted non-exemplars. Chanele's average accuracy on tact maintenance blocks was only 83.3% for category exemplars and 79.9% for the entire block. However, prior to listener and

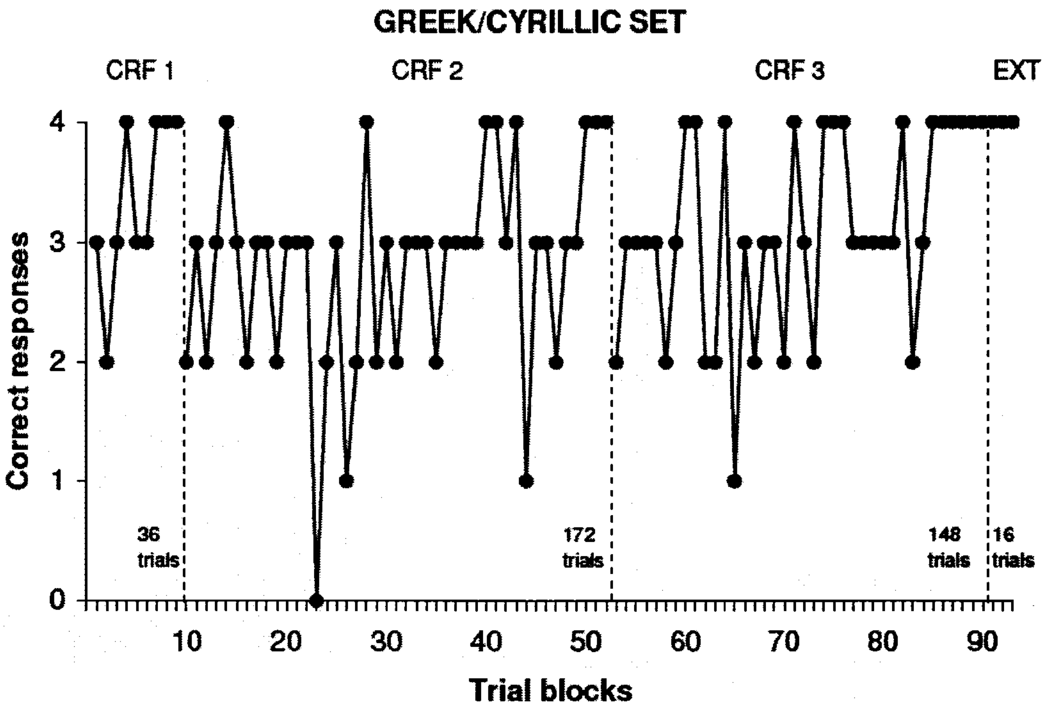
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

intraverbal categorization post-testing, she had responded with 100% accuracy on the first block in the last three sessions in which tact maintenance trials were conducted. Similarly, Sam's average accuracy for category exemplars in tact maintenance was 87.5%; however, his performance was highly accurate at the time of post-testing, and deteriorated only when retraining on the intraverbal repertoire began. Block-by-block data for tact and listener training and maintenance for each child are shown in Appendix C, along with data on interobserver agreement in those conditions.

Categorization Training and Testing

Progress of training. Figures 3, 4, and 5 depict intraverbal categorization training data for Marc, Greg, and Sam, respectively, along with the number of trials to criterion in each stage of training. Marc and Greg received training on two category sets each, and Sam was trained on one category set. On Figures 3 and 4, the upper panel shows training data for the first category set trained, whereas the lower panel depicts data for the second set. Marc reached the training criterion on the Greek/Cyrillic and East/West sets in 372 and 544 trials, respectively. Greg reached criterion in 688 trials on the North/South category, but an additional 68 trials of retraining under EXT were later required because of declining performance during post-testing. On the Kata/Hira category set, Greg reached the final training criterion in 1268 trials. Performance on that set declined during the first EXT condition, as a result of which the CRF condition was reinstated, followed by a second round of EXT. Finally, Sam reached criterion on the Greek/Cyrillic set in 420 trials. As with Greg, retraining on this set was necessary for Sam; however, Sam's participation in the experiment was discontinued before retraining was completed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



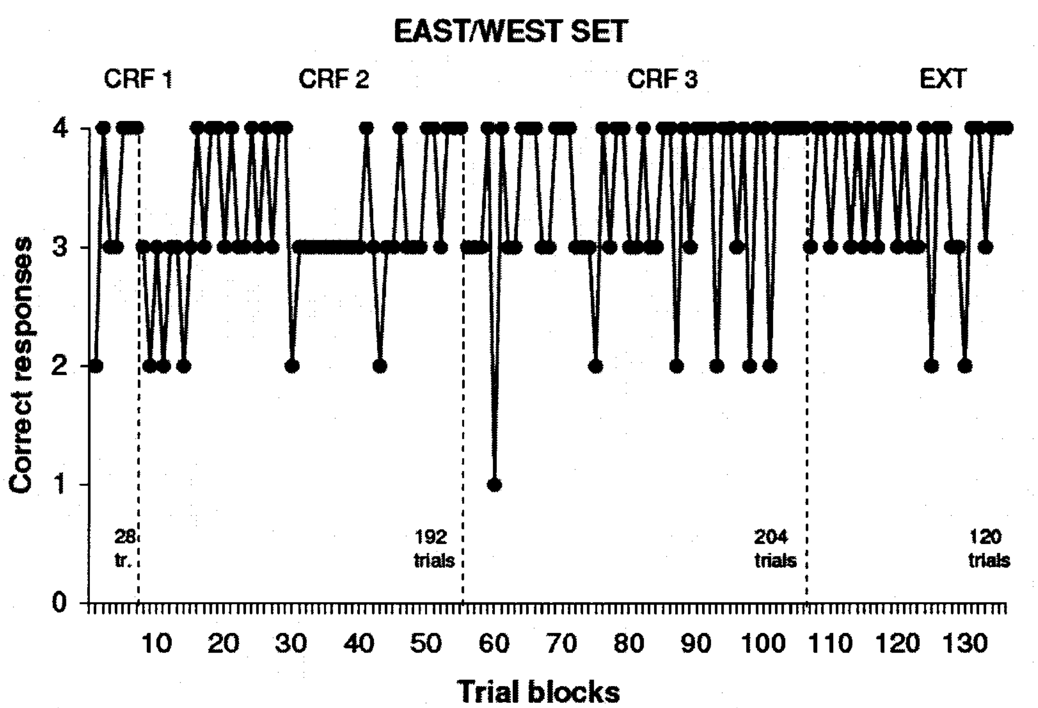
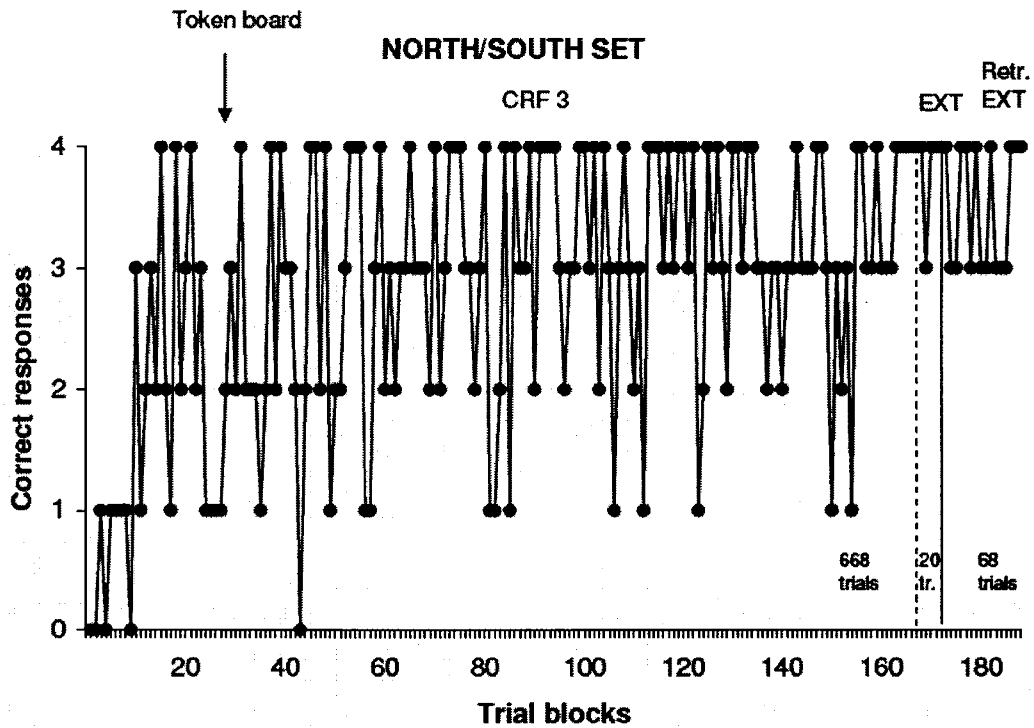


Figure 3. Frequency of correct responses across trial blocks during intraverbal categorization training for Marc.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



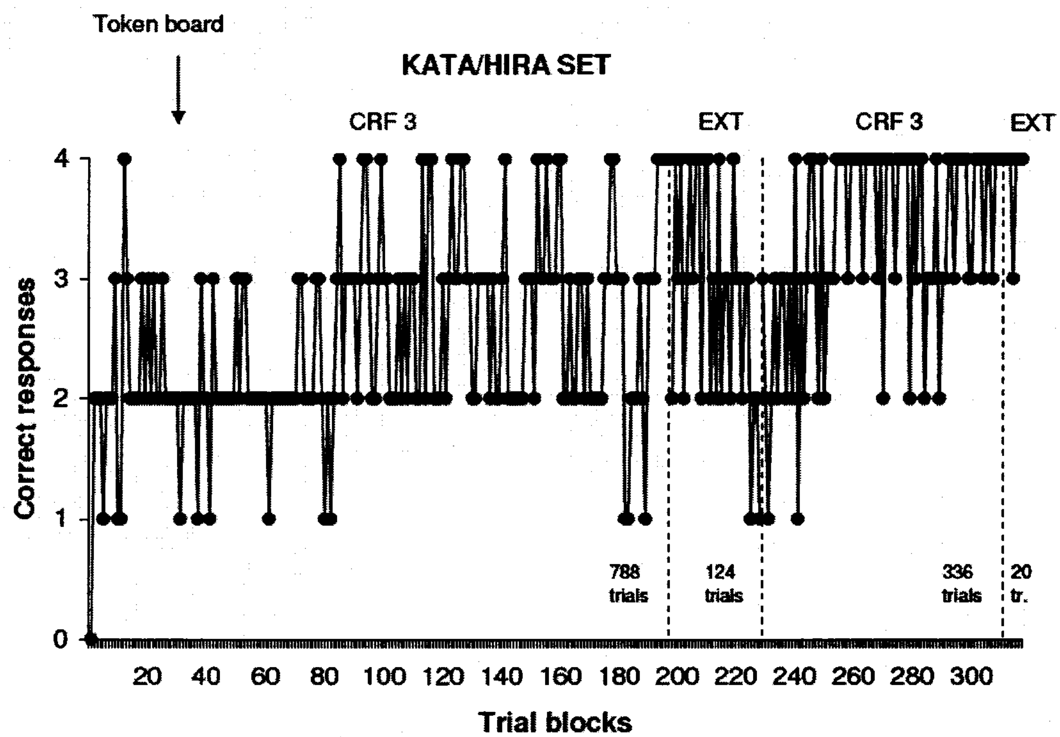


Figure 4. Frequency of correct responses across trial blocks during intraverbal categorization training for Greg.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

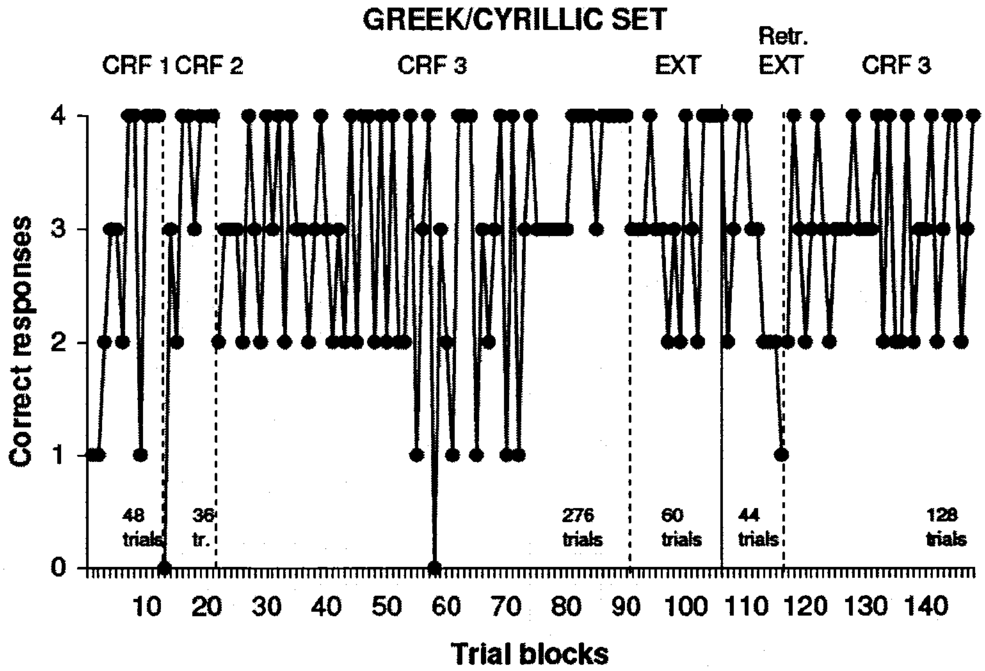


Figure 5. Frequency of correct responses across trial blocks during intraverbal categorization training for Sam.

Figure 6 shows listener categorization training data for Erika, and Figure 7 shows data for Chanele and Sanjay. For Erika, training on the East/West set (upper panel) and the Greek/Cyrillic set (lower panel) was accomplished in 804 and 340 trials, respectively. Chanele completed training on the Kata/Hira set in 932 trials, and Sanjay on the North/South set in 212 trials.

The average number of trials to criterion was 658.4 in intraverbal training, and 576.0 in listener training. However, intraverbal training did not consistently take longer to complete than listener training. Overall, training took longest to complete for Greg and Chanele, whose training procedures differed slightly from those employed for the other children.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

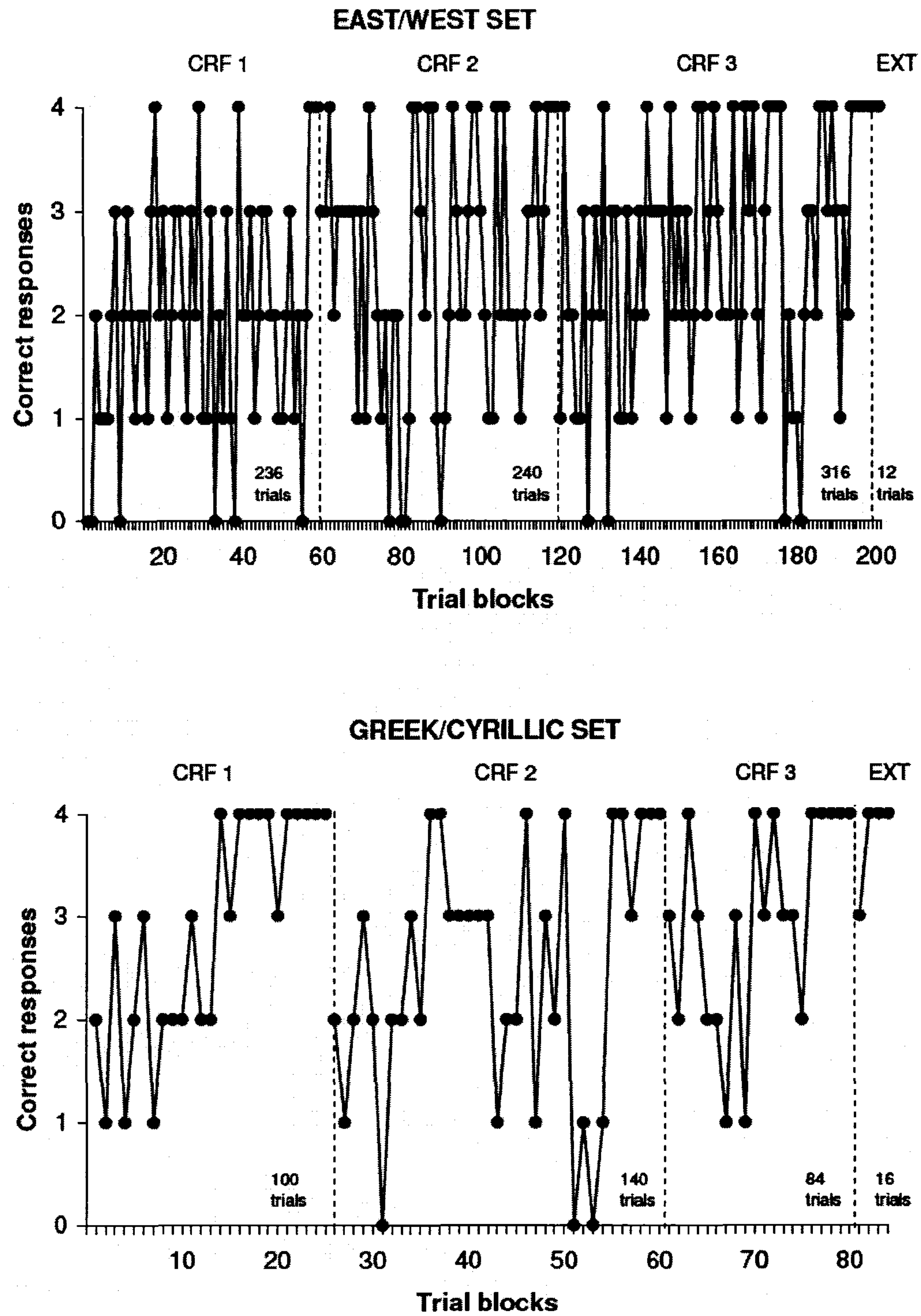
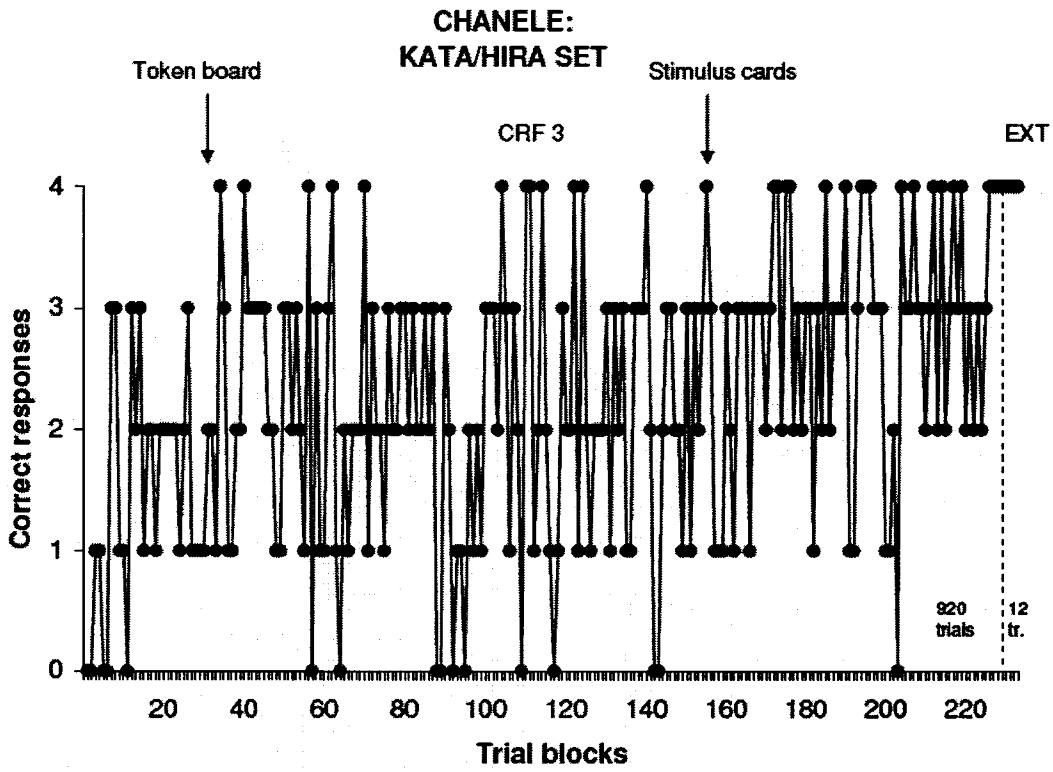


Figure 6. Frequency of correct responses across trial blocks during listener categorization training for Erika.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



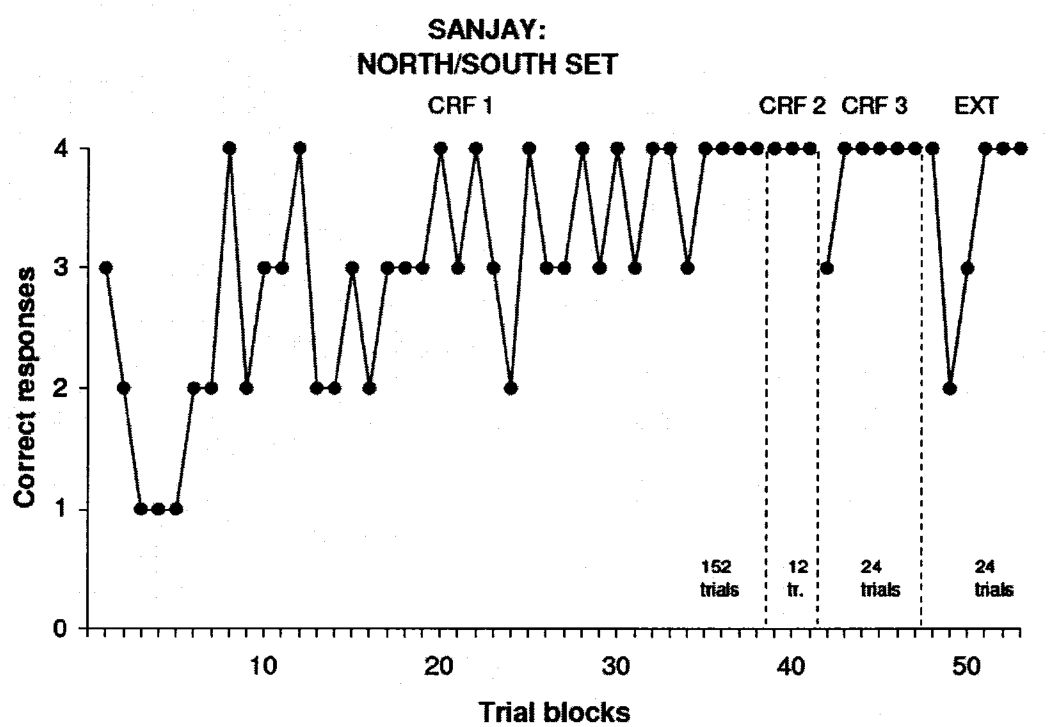


Figure 7. Frequency of correct responses across trial blocks during listener categorization training for Chanele and Sanjay.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Effects of intraverbal categorization training. Figures 8, 9, and 10 show the effects of intraverbal training on intraverbal and listener categorization for Marc, Greg, and Sam, respectively. On Figures 8 and 9, the upper panel shows data on the first category set trained, while the lower panel shows data on the second set. Sam received training on only one set, shown in the upper panel of Figure 10, while the lower panel depicts baseline data on a second set. The shaded portion of each figure contains data collected on the second category set during training on the first set. The horizontal dashed lines on each figure indicate chance level responding on listener categorization trials.

In baseline on the Greek/Cyrillic category set, Marc emitted the responses "Greek" and "Cyrillic" in most trials and thus his baseline performance was around 50% accurate (see Figure 8). Due to experimental error, Marc's baseline testing on the East/West category set did not begin until after 11 blocks (44 trials) of intraverbal categorization training on the East/West category set. Initially, no correct intraverbal responding occurred during baseline testing on the East/West set; however, following completion of training on the Greek/Cyrillic set, he began to consistently respond with 50% accuracy in the East/West baseline; responding with "East" in most trials. Although all of the children were exposed to category names in baseline trials for listener categorization, Marc was the only one who reliably emitted intraverbal responses in baseline. In listener trials, Marc's responding in both baselines was overall around chance level.

Following intraverbal training on the Greek/Cyrillic set, Marc's intraverbal responding was 100% accurate on 6 out of 10 testing blocks. Correct responding in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

listener trials also increased from baseline. During the first nine blocks of testing, he responded correctly on 3 to 4 trials in each block. This was, however, not sufficient to meet the acquisition criterion of three consecutive blocks of 100% correct performance. The effect on listener responding was not replicated with the East/West set, as following training on that set, he continued to respond at chance levels in listener trials. Performance in intraverbal trials, however, remained highly accurate.

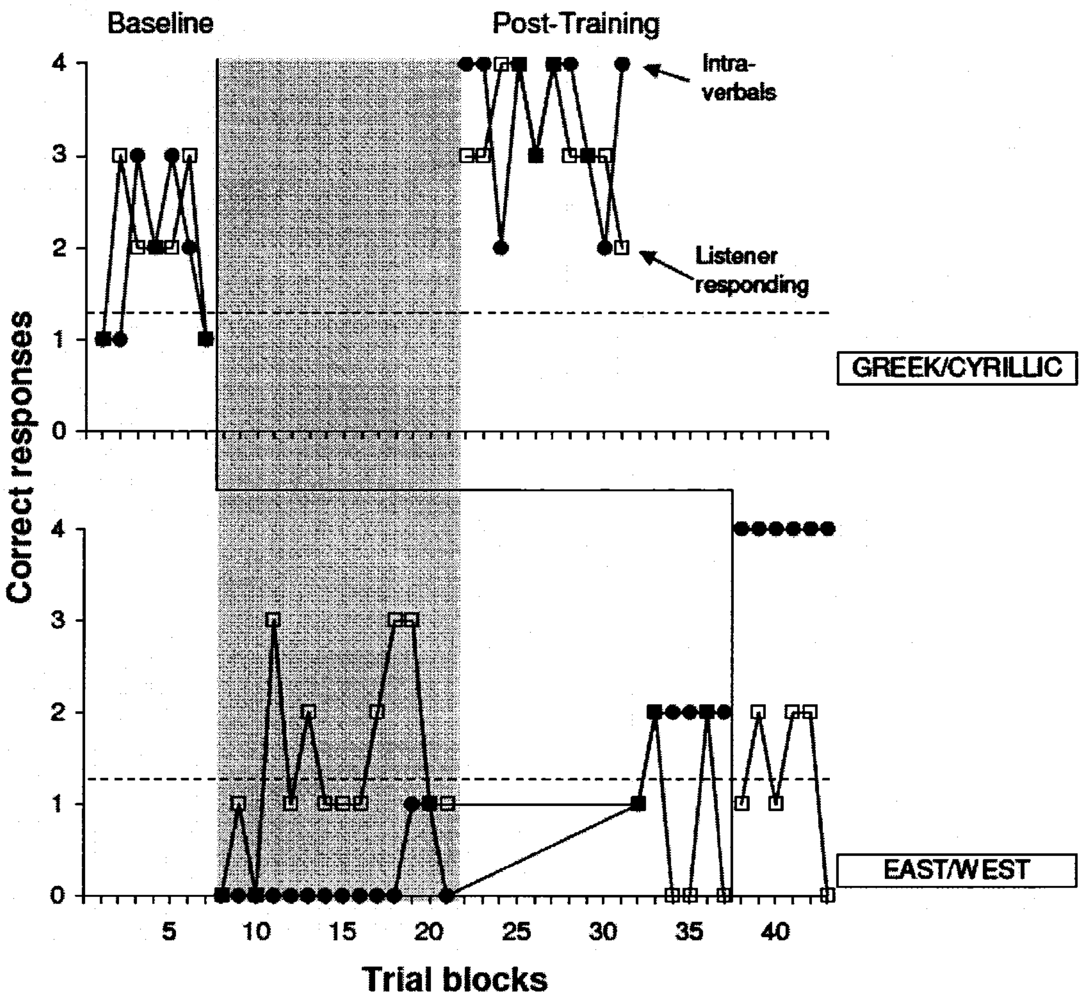


Figure 8. Marc's performance on baseline and post-training tests of listener and intraverbal categorization.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

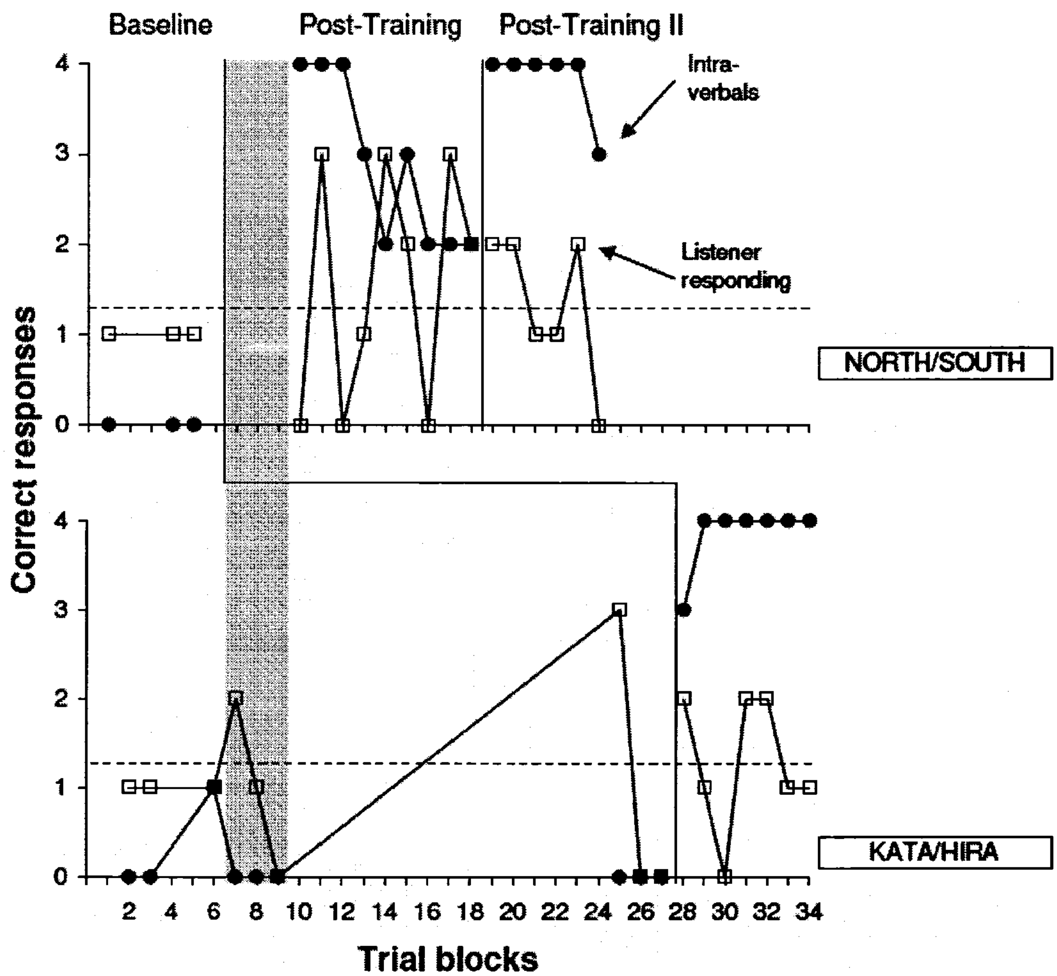


Figure 9. Greg's performance on baseline and post-training tests of listener and intraverbal categorization.

Greg responded around chance level in listener trials in baseline, and emitted only one correct intraverbal response, in the Kata/Hira baseline (Figure 9). During post-testing on the North/South category set, Greg's performance in intraverbal trials declined. Following retraining, his intraverbal performance remained accurate; however, performance in listener categorization trials did not exceed chance level. Following training on the Kata/Hira category set, Greg's performance in listener trials

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

also remained at chance level, while he responded with high accuracy in intraverbal trials.

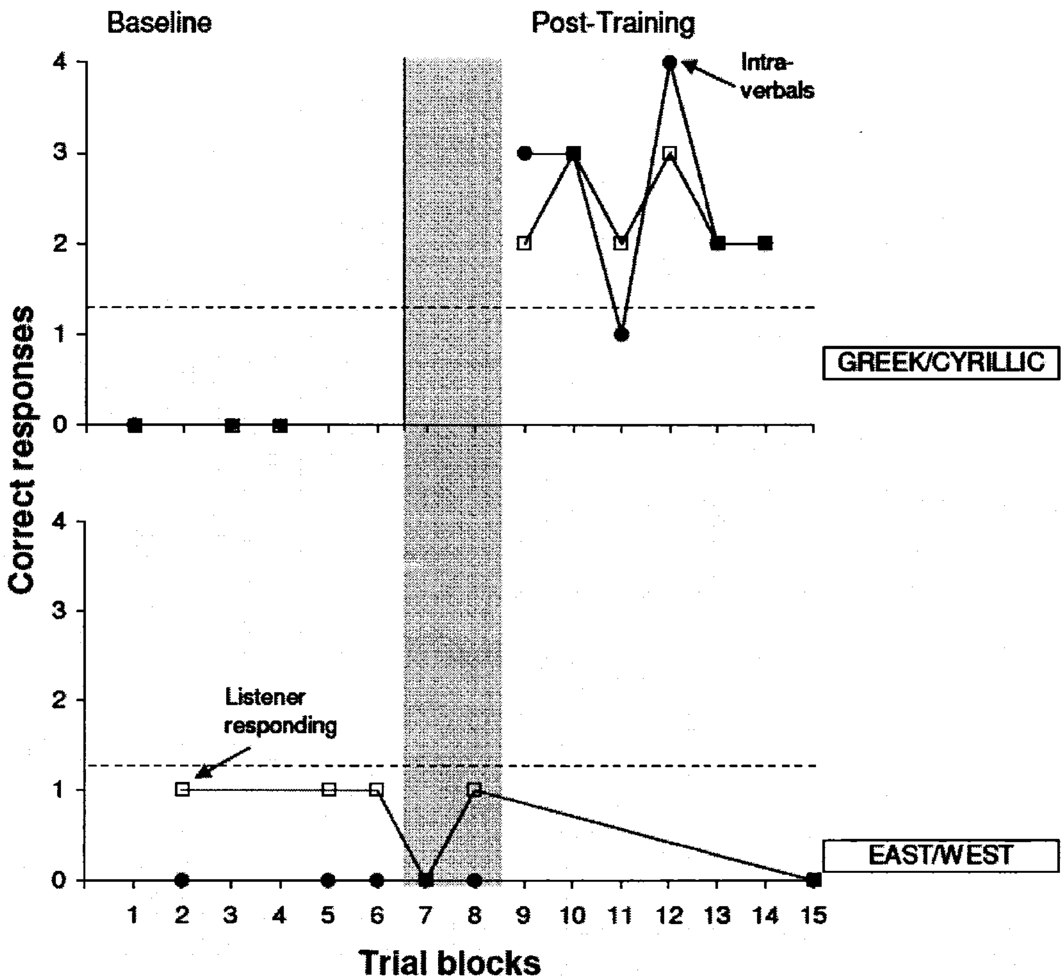


Figure 10. Sam's performance on baseline and post-training tests of listener and intraverbal categorization.

As shown in Figure 10, Sam made no correct intraverbal responses in baseline. In baseline listener trials, he frequently failed to respond at all, as a result of which he made no correct responses in the Greek/Cyrillic baseline, but performance in the East/West baseline was around chance level. Following intraverbal categorization training on the Greek/Cyrillic set, Sam's listener performance consistently rose above

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

chance level, as he responded correctly on 2 to 3 trials in each block. The accuracy of intraverbal responding did not maintain during post-testing, and therefore an attempt was made to retrain the intraverbal categorization repertoire. Before training was completed, however, Sam transferred to another child-care center, preventing further evaluation.

Marc, Greg, and Sam's performance on pre- and post-tests of category matching, category tacts, and reverse intraverbal categorization is shown in Figure 11. The horizontal dashed line on each graph denotes chance level on the category matching test (33.3%). For Greg and Sam, no changes were observed from pre- to post-test, except that on the category tact post-test following his first round of intraverbal training on the North/South set, Greg responded with "South" in most trials, resulting in 40% accurate performance. Although Greg's category matching performance was also higher on post- than pre-tests, it never exceeded 50%, and other children's pre-test data indicated that such performances could arise without training. Marc responded with 80% accuracy on the category tact post-test on the Greek/Cyrillic set, and with 70% accuracy on the reverse intraverbal categorization post-test on the same set. As with his performance on listener categorization trials, this increase was not replicated with the East/West set. No increase was observed from pre- to post-test in category matching on either set.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

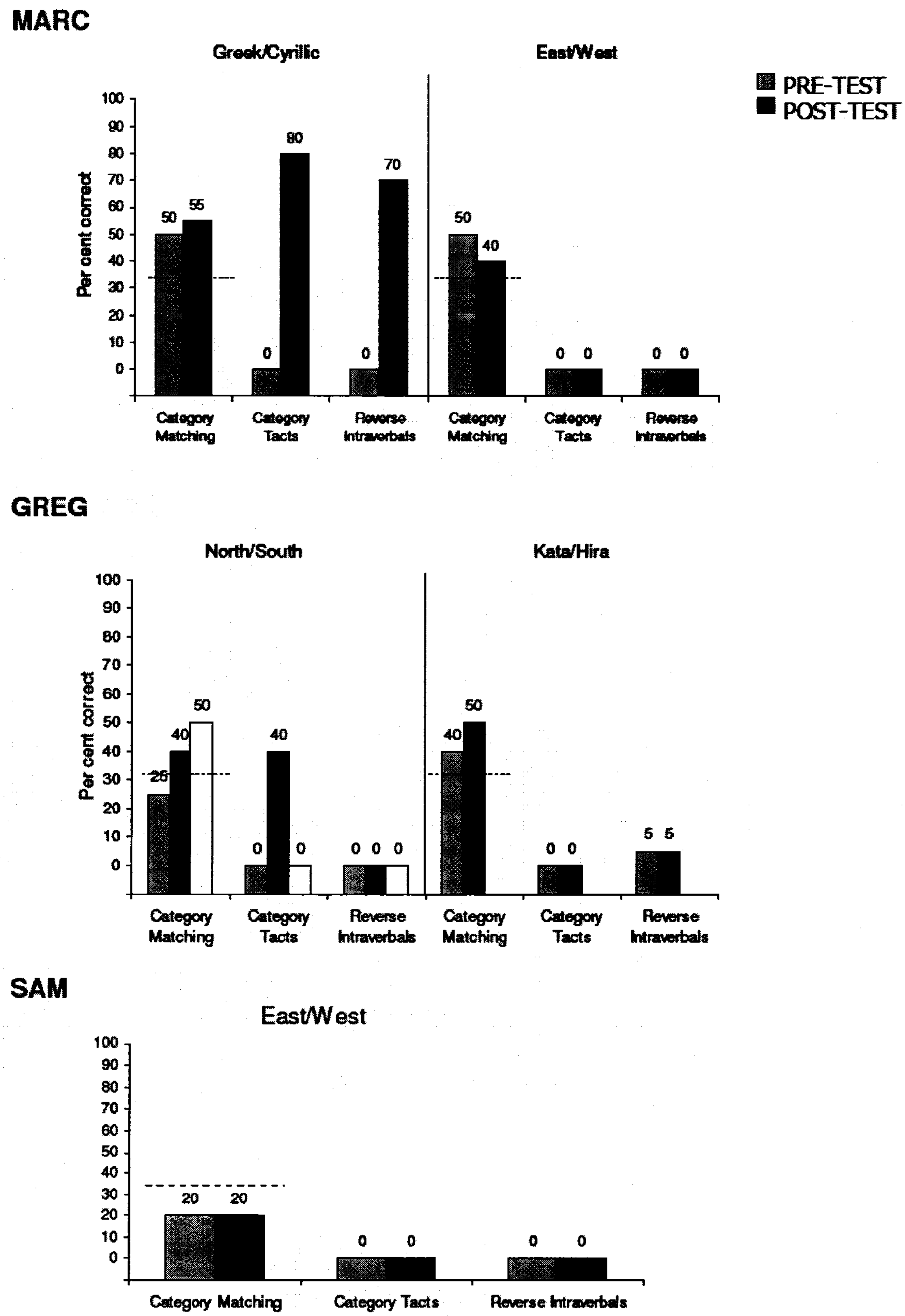


Figure 11. Marc, Greg, and Sam's performance on category matching, category tact, and reverse intraverbal pre- and post-tests.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Effects of listener categorization training. Figure 12 shows the effects of listener categorization training on listener and intraverbal categorization for Erika, and Figure 13 shows similar data for Chanele and Sanjay (baseline measures showing no change in Chanele's intraverbal or listener responding on an untrained category set are omitted from the graph). As in figures 8, 9, and 10, the horizontal dashed lines indicate chance level responding in listener trials.

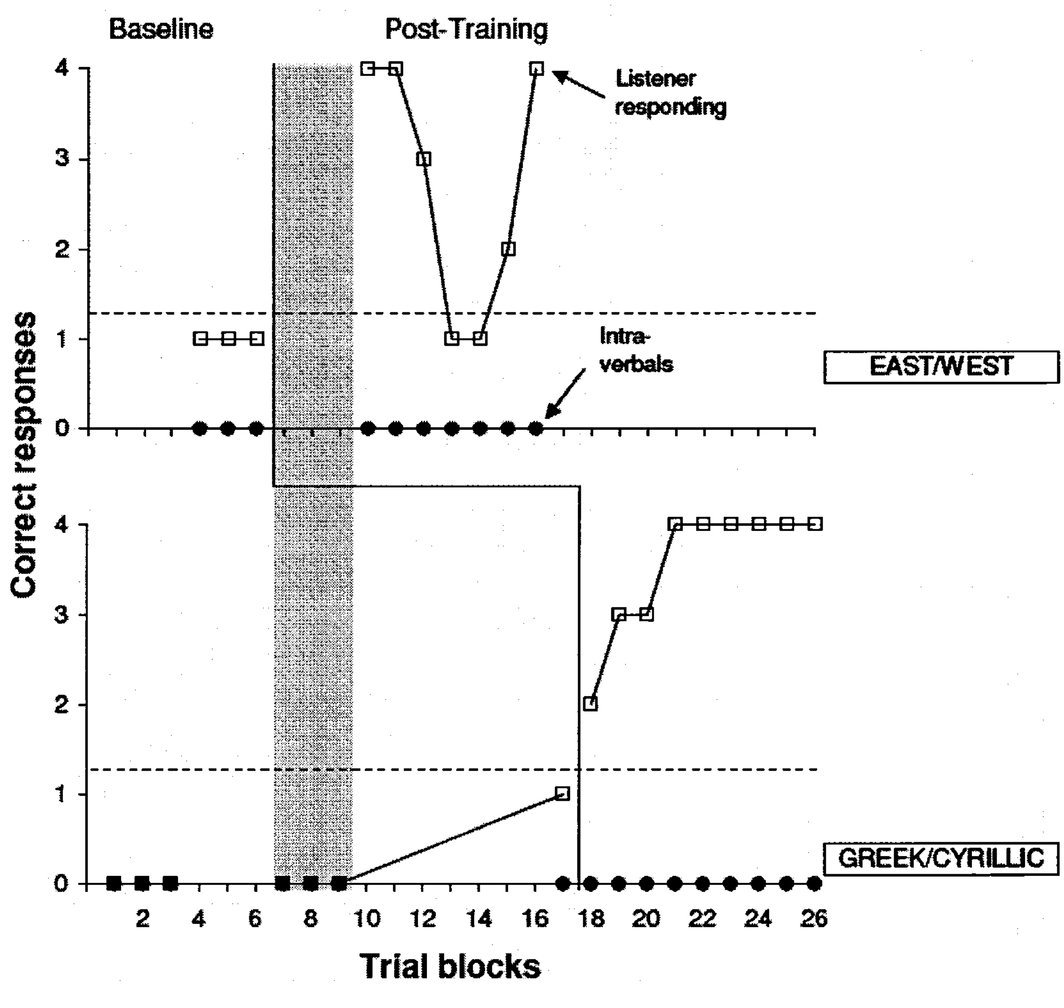


Figure 12. Erika's performance on baseline and post-training tests of listener and intraverbal categorization.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

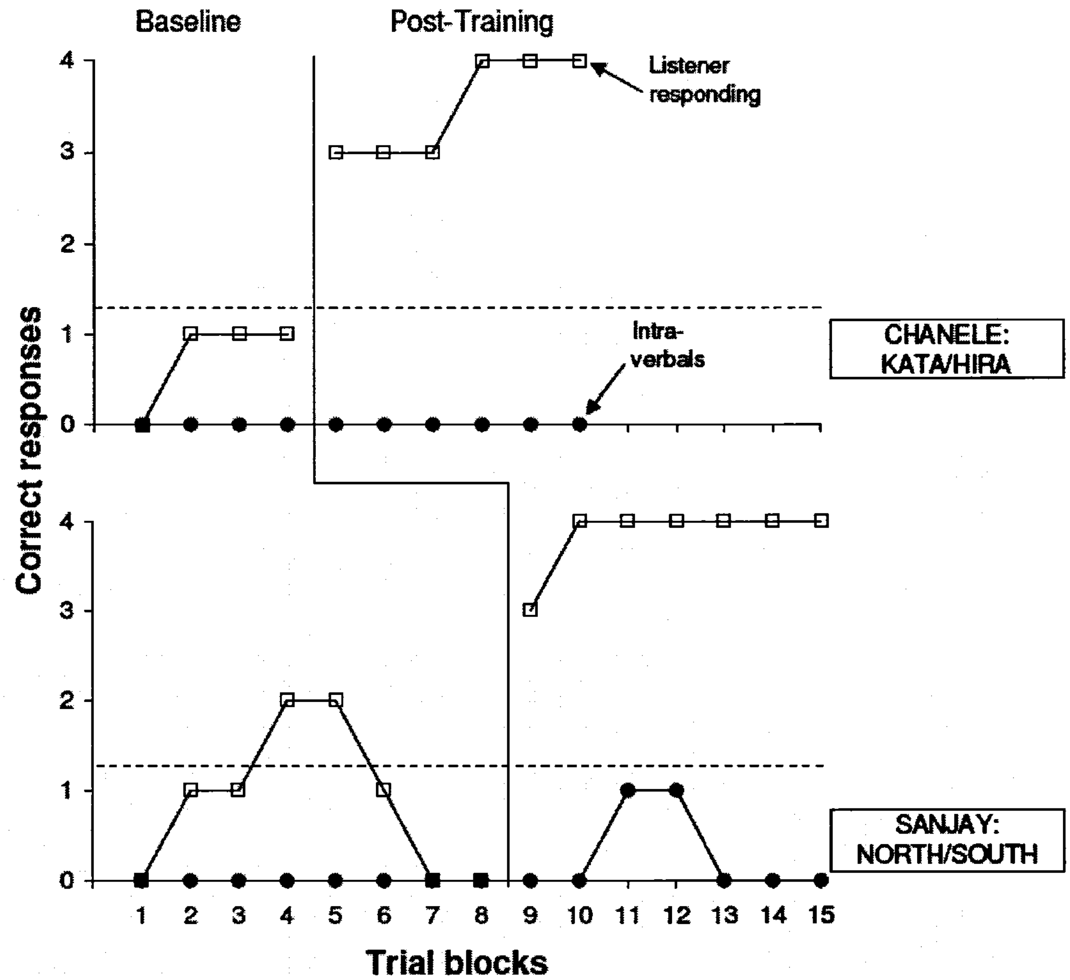


Figure 13. Chanele and Sanjay's performance on baseline and post-training tests of listener and intraverbal categorization.

Erika made no correct intraverbal responses in baseline on either the East/West or the Greek/Cyrillic set (see Figure 12), and performance in listener categorization trials was below chance level. Following listener categorization training on the East/West set, Erika's performance in listener categorization trials deteriorated, but accurate performance was recovered in the last block of testing. Following training on the Greek/Cyrillic set, accurate listener performance was observed beginning with the fourth block of testing. In neither case, however, did Erika emit any correct intraverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

responses following listener training. Similar results were obtained for Chanele and Sanjay (see Figure 13). Neither child responded intraverbally in baseline, while listener responding was around chance level. Following listener categorization training, both children's performance in listener trials was highly accurate, but no intraverbal responding emerged, with the exception that Sanjay emitted one correct intraverbal response in the third and fourth blocks of testing following training.

Figure 14 shows Erika, Chanele, and Sanjay's pre- and post-test performance in category matching, category tacts, and reverse intraverbal categorization. For Erika and Chanele, no increase in category matching was observed from pre- to post-test, and no accurate category tacts or reverse intraverbals were emitted with the exception of one category tact for Chanele. Sanjay's performance did not improve substantially on any of the tests; however, he responded correctly on a few category tact and a few reverse intraverbal categorization trials on the post-test.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

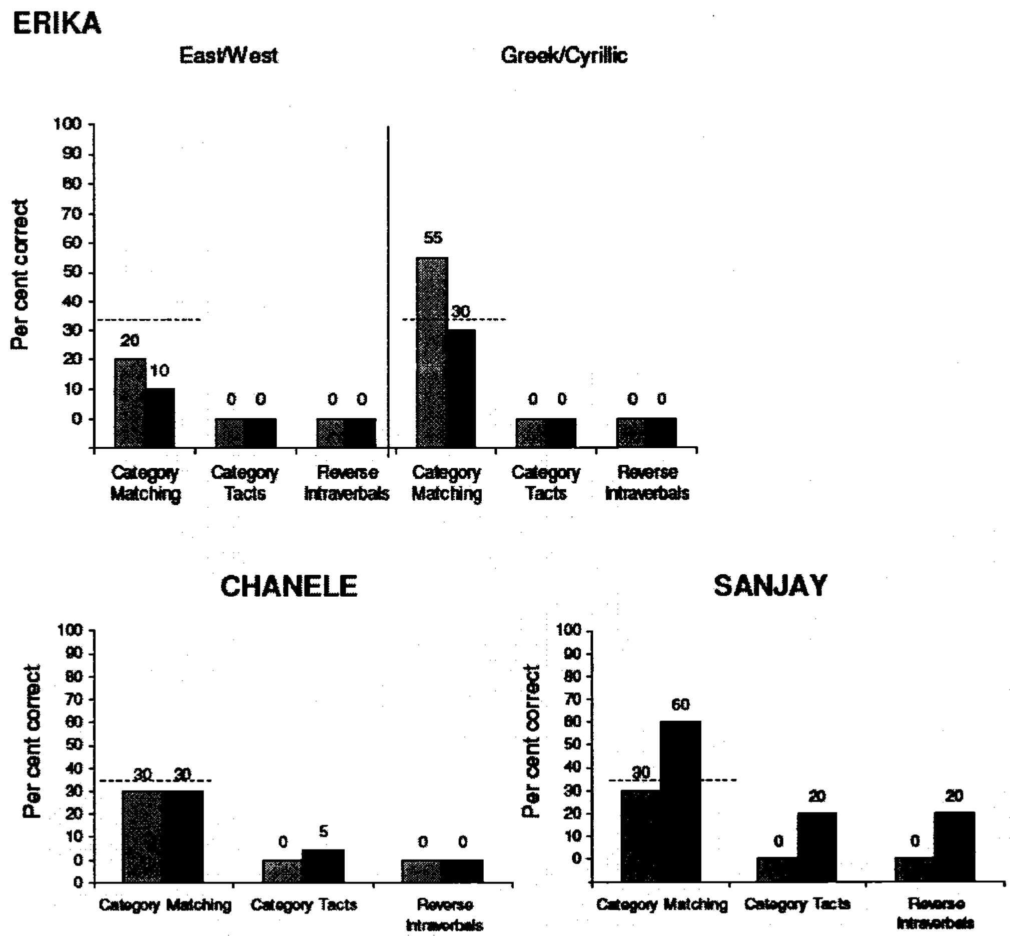


Figure 14. Erika, Chanele, and Sanjay's performance on category matching, category tact, and reverse intraverbal pre- and post-tests.

Collateral Verbal Responses

For most of the children, collateral verbal responses were rarely observed during listener and intraverbal categorization training and testing. The only exception was Sam, for whom collateral responses were recorded on 162 trials. A total of 155 of those responses occurred in intraverbal categorization training trials, and all consisted of echoic responses, such that instead of responding only with the category name as required, Sam would also repeat the experimenter's instruction and respond, for example, with "Psi is Greek", instead of just "Greek". This type of responding occurred

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in 36.9% of all intraverbal categorization training trials. During listener and intraverbal categorization post-testing, this type of responding was never observed on intraverbal trials. Two other collateral responses, however, occurred on listener trials during post-testing. Both consisted of echoing the category name contained in the experimenter's instruction prior to making a selection response, but in neither case did those echoic responses coincide with a correct selection response. The remaining five responses occurred during baseline listener and intraverbal categorization trials; in four cases Sam tacted the stimulus he selected in a listener trial, and in one case he echoed the exemplar name contained in the instruction in an intraverbal baseline trial.

Greg made 10 collateral responses, all of which occurred in listener categorization trials during post-testing, nine during post-testing on the North/South set, and one on the Kata/Hira set. In all cases he tacted the stimuli that he selected, but only in three cases did the tact coincide with a correct selection response.

Marc made only one collateral response, which occurred on a listener categorization trial during post-testing on the Greek/Cyrillic category; however, it was recorded by only the primary but not the secondary observer. According to this observer, Marc emitted both an echoic response and a tact on this trial; that is, he said "Greek is Psi", and then selected the correct comparison stimulus, Psi.

The five collateral responses recorded for Erica all occurred during listener categorization training. Two were echoic responses, and two were tacts of the selected stimulus. Chanele was observed to make 16 collateral responses, all during listener categorization training and listener and intraverbal categorization post-testing. During training, Chanele echoed the category name on four trials, tacted the selected stimulus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on three trials, and on one trial she emitted both an echoic response and a tact, saying "the mo is the hira", followed by a correct selection of mo. During post-testing, collateral responses occurred on seven listener trials. In five cases, Chanele echoed the category name, and in two cases she tacted the selected stimulus, which in both cases was the positive comparison. One echoic response occurred on an intraverbal trial, and was not followed by a correct intraverbal response. Finally, no collateral responses were recorded for Sanjay during either training or testing.

DISCUSSION

Summary of Results

No convincing evidence emerged that in the context of teaching categorization skills to young children, listener training might result in the emergence of an intraverbal repertoire or intraverbal training in the emergence of a listener repertoire. The three children who received listener categorization training very rarely emitted correct intraverbal responses when tested following training, and this finding was replicated with a second stimulus set for one of the children. Similarly, for one of the three children who received intraverbal training, no increase in listener responding was observed. For another child, Marc, an increase in correct listener responding was observed following intraverbal training on the first set; however, he did not meet the acquisition criterion for this untrained repertoire. Further, no such increase was observed following intraverbal training on the second set, and thus experimental control was not achieved over this effect. For the third child, Sam, a slight increase in accuracy appeared in listener trials following intraverbal categorization training. However,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

because the accuracy of Sam's intraverbal responding deteriorated rapidly during testing, and his participation in the study was discontinued before it could be retrained, it is unknown whether he would have met the acquisition criterion for the listener repertoire if his intraverbal repertoire had maintained.

Further, there was little evidence that either type of training produced any of the relations that were defined for the present purposes as reverse intraverbals, category tacts, and category matching. The only exception occurred for Marc, for whom intraverbal training on the first set appeared to generate some category tacts and reverse intraverbals, and this effect coincided with the increase in correct listener responding observed following intraverbal training on the same set.

Functional Independence of Intraverbal and Listener Behavior

The results of the present study are consistent with the notion of the functional independence of speaker and listener behavior suggested by Skinner's (1957) analysis of verbal behavior, in that listener training did not generate intraverbal responding, and intraverbal training tended not to generate listener relations. The failure of either type of training to generate untrained listener or intraverbal relations cannot be explained by a failure of training effects to maintain during testing. The previously trained tacts and listener relations, which could have potentially served to link listener and intraverbal relations, were maintained at high accuracy throughout the study. Further, performance on trials that tested directly trained listener and intraverbal relations remained accurate during the testing of untrained relations, except in the case of Sam, who did not complete the study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The data on the effects of listener categorization training on intraverbal responding are consistent with the findings of Miguel et al. (2005), and extend it in several ways. First, in the Miguel et al. study, each experimentally defined category contained 10 exemplars. During each instance of listener categorization training, the participants were trained to select each of the 10 visual stimuli that belonged to one category given its spoken category name, as well as its exemplar name. The dependent variable was performance on intraverbal probes on which a spoken category name was presented, and the participants' task was to respond vocally with exemplar names. Thus, up to 10 correct responses were possible on a given probe. Although listener training in no case produced large or consistent increases in intraverbal responding, some of the participants did emit intraverbal responses on at least some probes following training, which brings up the possibility that more accurate performances might have been observed had the experimentally defined stimulus classes been smaller. In the present study, the participants received training on two categories at a time, and each category contained only two exemplars. Thus, in each reverse intraverbal categorization trial, only two correct responses to a spoken category name were possible. Yet, little or no increase in reverse intraverbal categorization was observed following listener categorization training. Second, Miguel et al. did not assess the effects of listener training on intraverbal responding in which the exemplar name served as the S^D and the category name as the correct response, which was the primary dependent variable in the present study. The present data thus extend Miguel et al.'s findings (a) to the effects of listener training on an intraverbal relation in which there is only one correct response to each verbal stimulus, but the same response is correct in the presence of more than one

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stimulus, and (b) to the effects of listener training on an intraverbal relation that does not share an S^D with the trained listener relation. Listener training did not establish this type of intraverbal responding in any case. Third, Miguel et al. pointed out that different results might be obtained if more than one category at a time were targeted for training, similar to research on stimulus equivalence, in which participants typically receive training on at least two experimentally defined stimulus classes at a time. In the present study, two categories were trained at a time, and yet no intraverbal responding emerged. Finally, it did not appear to make a difference that the participants in the present study were initially trained to tact all category exemplars (which was not the case for those children who received listener training first in Miguel et al's study), as the topographies trained as tacts did not appear on the reverse intraverbal categorization test. The two studies thus appear to suggest that listener training may not easily establish intraverbal relations among young typically developing children, at least in the context of teaching categorization skills.

The present data on the effects of intraverbal categorization training differ from those of Luciano (1986), C. T. Sundberg and Sundberg (1990), and Remington and Clarke (1993b), all of whom found that intraverbal training resulted in the emergence of listener responding. The intraverbal relations trained in the present study, however, differed from those trained in the prior studies in more than one way. For example, in two of the prior studies (Remington & Clarke, 1993b; C. T. Sundberg & Sundberg), intraverbal training involved the training of one intraverbal response to each verbal stimulus. In the present study, by contrast, the same intraverbal response was established in the presence of two different verbal stimuli. Possibly this more complex

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

type of training is less likely to establish listener relations. Further, the intraverbal relations trained in those same studies were signed responses to spoken verbal stimuli, whereas in the present study they were vocal responses. Luciano, however, trained vocal intraverbals, and in her study, multiple intraverbal responses were established in the presence of the same verbal stimulus.

Another potentially important difference is that in all of the prior studies (Luciano, 1986; Remington & Clarke, 1993b; C. T. Sundberg & Sundberg, 1990), the intraverbal relation trained was one in which the verbal S^D was identical to the verbal stimulus presented when the listener relation was tested, and the response topographies had previously been trained as tacts of the visual stimuli involved in the listener relations. This was not the case in the present study, in which the verbal S^D for the trained intraverbal corresponded to the verbal S^D for an already trained listener relation, but differed from the verbal stimulus presented when listener behavior was tested. The stimulus presented on listener trials instead corresponded to the product of the trained intraverbal response. Further, the trained intraverbal response topography had not previously been trained as another verbal operant during the experiment, although it had been established that the participants were able to emit it as an echoic response. Thus, in the present study, intraverbal training did not directly establish the verbal and the nonverbal stimulus as functionally equivalent stimuli both evoking the same vocal response topography. Some contemporary behavior-analytic accounts of language appear to suggest that this may be of relevance. For example, Lowenkron's (1998) joint control analysis appears to suggest that an untrained selection response, such as the listener responses tested in this study, may appear when two stimuli simultaneously

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

evoke the same verbal response topography, perhaps at a covert level. For example, the verbal stimulus (category name in the present study) and the nonverbal stimulus presented on a listener trial might evoke the same topography as an intraverbal and a tact, respectively. In the prior studies, such tacts and intraverbals had been directly established, whereas in the present study, only the tact had been directly established, as intraverbal responding was trained only to the exemplar name and not to the category name. Alternatively, it may be speculated that such a listener relation could emerge if the verbal stimulus (category name) evoked an intraverbal response (exemplar name), which in turn evoked the listener response of selecting the exemplar. Such an analysis would be in line with that suggested by Horne and Lowe (1996), although without assuming a higher-order naming relation. In the prior studies, the relevant intraverbal relation had been directly established, and in two of them, the listener relation was previously acquired (Luciano) or shown to have emerged as a result of other training (Remington and Clarke, 1993b), whereas its presence in the third is unknown (C. T. Sundberg & Sundberg). In the present study, only the listener relation had been directly trained, but again, the relevant intraverbal relation had not been trained. Both analyses, in other words, appear to suggest that intraverbal responding to the category name may be critical for a listener relation to emerge with respect to the category name. In the present study, this intraverbal relation was not directly established. Further, tests of reverse intraverbal categorization revealed that intraverbal training did not typically establish this relation. In fact, the only case in which there was evidence that the reverse intraverbal relation emerged was also the only case in which there was a clear increase

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in listener categorization. This effect occurred with Marc, who was additionally observed to overtly emit the reverse intraverbal relation on a listener test trial.

Sam's data, however, may not necessarily support this interpretation. Although his post-testing data are unclear due to the deterioration of intraverbal responding, it appears that he also may have been acquiring listener relations as a result of intraverbal training. Sam, however, did not respond correctly on a single test trial for reverse intraverbal responding. It is not surprising that his performance on this test was not highly accurate, because at the time of testing, stimulus control had already been lost over the trained intraverbal repertoire. However, if Sam had initially acquired the reverse intraverbals, perhaps loss of stimulus control should have been expected to result in around 50% correct responding rather than none at all. To the extent that Sam's listener performance did improve, therefore, it is conceivable that it did so in the absence of the reverse intraverbal relation. It may be argued that although an analysis in line with that of Horne and Lowe (1996) appears to require the reverse intraverbal relation for listener responding to emerge, a joint control analysis in line with Lowenkron's (1998) analysis does not necessarily require it. Specifically, if during testing, the children were to tact the comparisons with their exemplar names, this might evoke the already trained intraverbal category name response, simultaneously evoked by the instruction. Thus failure to respond correctly on listener trials might occur due to failure to tact the comparisons or respond echoically to the instruction. It should be noted, however, that across all participants, in the very few instances in which overt tacts or echoics were observed on listener test trials, they did not necessarily coincide with correct selections.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In sum, the present study along with prior research on the relation between listener behavior and topography-based intraverbal behavior may suggest that (a) the training of listener relations does not readily produce intraverbal relations via other established listener relations and tacts, and (b) the extent to which intraverbal training establishes new listener relations via already established tacts and listener relations may depend on the direction of the intraverbal relation trained. If this is the case, it may suggest that findings from the literature on stimulus equivalence and other derived selection-based relations are not directly applicable to the training of topography-based verbal relations. This literature, for example, would appear to predict the same results regardless of the direction of training of the intraverbal relation.

Alternative interpretations are possible. For example, in a conditional discrimination task, the presence of comparison stimuli has been said to provide a context for responding, whereas in topography-based relations, contextual stimuli other than the stimuli directly involved in the relations themselves may be necessary to evoke an appropriate response (D. Barnes-Holmes et al., 2000; Hall & Chase, 1991). In the present study, the experimental instructions should have served this function, and it is possible that the tested relations failed to emerge because the instruction failed to evoke appropriate responding that might have been evoked under other circumstances. In the pretraining condition, it was established that many of the experimental instructions already evoked appropriate responding to familiar stimuli, and appropriate responding in their presence was furthermore directly reinforced. Thus, it may seem implausible that they failed to evoke such responding during testing with the experimental stimuli. However, it is possible that the instructions were effective only in the presence of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stimuli with which the children were already familiar from their everyday environment, or with categories in which the exemplars shared physical features (such as animals) or functions (such as fruits). In addition, all of the children did require some amount of training with at least some of the instructions. It is conceivable that the training criterion was simply too lenient, as perhaps evidenced by the fact that additional training was sometimes required when the instructions were reviewed with familiar stimuli prior to testing with the experimental stimuli. Hence the possibility exists that a more extensive history of responding to the experimental instructions would have generated untrained listener and intraverbal relations. With respect to the implications for language training; however, this may not make a major difference. In EIBI, for example, when categorization programs are introduced, the learner with no prior categorization skills is almost unavoidably being exposed to the training instructions and context for the first time. Therefore, they should perhaps not be any more likely than children in the present study to acquire similar untrained relations.

Interestingly, although none of the participants demonstrated complete acquisition of untrained listener relations or verbal operants following either listener or intraverbal categorization training, five out of six children had previously acquired exemplar name listener relations as a result of exemplar name tact training. The sixth child required only minimal listener training following tact acquisition. These data are consistent with several prior studies which have demonstrated that young typically developing children may readily acquire appropriate listener relations as a result of tact training (e.g., Cuvo & Riva, 1980; Connell & McReynolds, 1981; Lipkens et al., 1993; Lowe et al., 2002, 2004). Thus, functional independence between speaker and listener

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

behavior was observed when the speaker relation in question was an intraverbal, but functional interdependence was observed when it was a tact. As previously noted, contemporary behavior-analytic accounts of language acquisition may suggest either that functional interdependence of speaker and listener behavior arises via unobserved occurrence of collateral responding (e.g., Horne & Lowe, 1996) or that it is an instance of relational responding attributable to a history of multiple-exemplar training in an appropriate context (D. Barnes-Holmes et al., 2000). Both types of analyses appear to suggest that functional interdependence between listener relations and tacts may be likely to arise earlier in a developmental sequence than functional interdependence between listener relations and intraverbals. Verbal operant analysis in line with those offered by Horne and Lowe, or Lowenkron (1998), suggest that more complex collateral responding is required for intraverbal training than tact training to generate listener relations. Similarly, an analysis in terms of derived stimulus relations would suggest that the emergence of listener relations as a result of tact training is analogous to symmetry; a necessary but not a sufficient condition for transitivity, which would be analogous to the emergence of listener relations as a result of intraverbal training. Future research might investigate the extent to which in typical language acquisition, the emergence of functional interdependence of verbal operants and listener relations occurs in a predictable sequence.

Effects of Training on Category Matching

The category matching post-tests revealed that none of the participants were able to match visual stimuli to one another following either listener or intraverbal training. It is somewhat surprising that listener training did not produce category

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

matching, given that a similar type of auditory-visual training has tended to reliably give rise to similar visual-visual matching performances in the stimulus equivalence literature. However, the present study differed procedurally in several respects from typical stimulus equivalence protocols. For example, in stimulus equivalence research, test trials for emergent relations are often interspersed with trials that test the trained baseline relations, and correct responses on those trials are sometimes reinforced (Green & Saunders, 1998). Thus, it is verified that the baseline relations maintain during testing. In the present study, although the trained listener relations typically maintained throughout the testing of intraverbal categorization, no listener trials were conducted during category matching, category tact, or reverse intraverbal categorization post-tests, and it is possible that the listener relations deteriorated during this time. Further, it has been common practice by some stimulus equivalence researchers, if untrained relations do not emerge when initially tested, to retrain baseline relations and/or expose the participant to additional testing of untrained relations (Green & Saunders). No such procedures were employed in the present study and, thus, it is unknown whether category matching would have emerged had testing continued longer, with or without intervening retraining of listener relations. In fact, the same can be said of all other relations tested in this study.

It also may be worth noting that the children in this study were all three years of age at the beginning of the experiment. Although young children have frequently participated in research on stimulus equivalence, relatively little research has been conducted with children under five years of age (see Boelens, Van den Broek, & Van Klarenbosch, 2000). When younger children have been studied, difficulties have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sometimes been reported with establishing the required baseline conditional discriminations (Augustson & Dougher, 1991; Pilgrim, Jackson, & Galizio, 2000). When conditional discrimination training has been successful, equivalence relations or similar performances have been shown to emerge in children as young as 2 or 3 years of age (e.g., Barnes et al., 1995; Horne et al., 2004, 2006; Jordan, Pilgrim, & Galizio, 2001). However, in Horne et al.'s (2004, 2006) studies, such performances were strongly correlated with the extent to which listener training (i.e., training of baseline relations) successfully produced tacts. When listener relations were not accompanied by tacts, accurate stimulus sorting, which was the primary dependent variable, did not emerge. Specifically, having learned to select two different stimuli in the presence of the same verbal stimulus, as did the listener-trained children in the present study, the children who failed did not tact the stimuli with a common topography, similar to category tacts tested in the present study. In the present study, none of the children who received listener training with respect to category names acquired the category tact repertoire as a result, and none of them acquired the category matching repertoire.

Intraverbal training also did not generate accurate category matching in the present study, even in the one case in which there was evidence that all other trained relations emerged. While the effects of topography-based intraverbal training on such a performance may not have been investigated before; a stimulus equivalence analysis would appear to predict its emergence. Empirically, however, transitive relations have been obtained less readily when the number of *nodes*, or stimuli that are related to more than one other stimulus, increases (Fields, Adams, Verhave, & Newman, 1990; Kennedy, Itkonen, & Lindquist, 1994), and in the present study, category matching was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

nodally more distant from trained relations in the intraverbal than in the listener training condition. Thus, given that not even listener training established category matching with the procedures used in the present study, intraverbal training perhaps should not have been expected to do so either.

It may be of interest to note that Horne and Lowe (1996) have hypothesized that performance on equivalence tests, for example, visual-visual matching following listener training, may be verbally mediated in at least two ways. First, training may produce common tact and listener relations (i.e., name relations) with respect to visual stimuli that belong to the same experimentally defined class, such that during testing, the participant may tact the sample and respond as a listener by selecting the appropriate comparison. Second, the participant may during training acquire idiosyncratic tacts and listener relations (or name relations) with respect to individual stimulus exemplars, which may then become linked intraverbally to a common name (as a result of listener training), or linked intraverbally with one another (as a result of visual-visual match-to-sample training). During testing, the participant's tact of the sample may thus evoke an intraverbal response or a chain of intraverbal responses, which in turn evokes a correct selection as a listener response. A joint control analysis (Lowenkron, 1998) similarly suggests that such test performances could result either from the presence of common tacts or of intraverbal relations. Stimulus equivalence researchers have frequently failed to find evidence that participants were able to emit common tacts of stimuli in the same class. Lowe et al. (2002, 2005), however, found that explicit training of such tacts, in the absence of other training, produced appropriate stimulus sorting, and as previously noted, Horne et al. (2004, 2006) found that the presence of such tacts was correlated

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with sorting performance. The possible role of intraverbal responding has received less attention, but it has been argued that it merits attention as well (e.g., Stromer, Mackay, & Remington, 1996). Recently, Smeets and Barnes-Holmes (2005) investigated and found no evidence that children in their study had acquired tacts that might permit equivalence class formation via intraverbal responding. In the present study, relevant intraverbal relations were trained directly, and yet no category matching emerged. Horne and Lowe's analysis, however, requires the intraverbal relations in question to be bidirectional, and in the present study, intraverbal training in most cases did not establish the reverse intraverbal relation. In the one case in which the emergence of reverse intraverbal responding was observed, category matching did not emerge; however, that participant's reverse intraverbal responding was not highly accurate.

Limitations

Several limitations of the present study may be noted. First, different children received listener and intraverbal training, and thus no within-subject evaluation was conducted of the effects of the two types of training. The small number of participants also does not allow for a between-subjects comparison of the effects of listener and intraverbal training. This would have been a major limitation if one type of training, but not the other, had been found to somewhat consistently produce the untrained relation. It then would have been unknown whether the results were due to the superiority of that type of training to the other, or due to different pre-experimental histories or other individual differences among children who happened to be exposed to different types of training. This was, however, not the case. Neither type of training reliably produced any untrained relations, and thus no questions arose over whether one type of training might

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

be more effective than the other in this regard. As a note of caution, however, due to the design of the study it is important not to interpret the limited evidence for Marc and Sam's acquisition of untrained relations as indicative that intraverbal training was more likely than listener training to produce such relations.

Second, experimental control was never achieved over the untrained repertoires. When untrained listener or intraverbal relations failed to emerge, no steps were taken to manipulate additional variables that might have established those relations, with or without direct reinforcement. It was therefore not demonstrated that children who received listener training would have acquired the tested intraverbal relations given appropriate environmental manipulations, or vice versa. However, although training was in some cases lengthy, all children eventually acquired the directly trained repertoires, whether they were assigned to listener or intraverbal training conditions. It therefore appears unlikely that children who did not acquire the intraverbal repertoire as a result of listener training would have been incapable of acquiring it at all, or vice versa, at least via direct training.

Third, all untrained relations were tested under extinction, which brings up the possibility that the children were not provided with sufficient motivation for correct responding on test trials. Greg and Sam's data indicate that at least in the case of intraverbal training, even the trained repertoire was sensitive to extinction. However, except in the case of Sam, who did not complete the experiment, it was eventually demonstrated that under extinction, the children continued to perform accurately on test trials for the trained repertoire. It therefore seems unlikely that the testing conditions can account for their poorer performance in intermixed trials that tested the untrained

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

repertoire. Nevertheless, it is possible that the untrained repertoire was somehow more fragile and more sensitive to the effects of extinction. In addition, the directly trained repertoire was not probed during post-tests of other categorization skills; that is, category matching, category tacts, and reverse intraverbals. As a result, it was not demonstrated that the trained relations maintained during those tests, nor that those testing conditions provided a sufficient level of motivation to respond correctly at all. Anecdotally, some of the children's short selection-response latencies on some test trials of listener categorization and category matching suggested that they might not be scanning the comparison stimuli. Further, on test trials that targeted untrained verbal operants, the children sometimes responded immediately with "I don't know.". Perhaps different results would have been obtained if the participants had been informed, for example, that correct responding on test trials would earn them tokens delivered at the end of the session (Galvao, Calcagno, & Sidman, 1992).

Finally, although this study arose out of considerations related to the sequencing of skills in behavioral language interventions, the children who participated were typically developing. As with all research, it is possible that different results would have been obtained had individuals representing other populations been studied. However, the literature on tact and listener training (see Tables 1 through 3) suggests that the training of verbal operants and listener relations may produce similar results for young typically developing children and individuals with diagnoses of developmental disabilities. To the extent that differences are observed, they appear to generally be in the direction of greater interdependence of speaker and listener repertoires among typically developing children. Since the children in the present study mostly performed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

poorly on all tests of untrained relations, it is therefore unlikely that different performances would be observed among children in need of language interventions.

Implications for Sequencing of Receptive and Expressive Programs

A primary purpose of the present study was to assess whether it might be feasible to train intraverbal relations before training listener relations in behavioral language interventions, as opposed to training listener relations first, as recommended in some EIBI curricula (Leaf & McEachin, 1999; Maurice et al., 1996). The present data mostly suggest that listener relations and intraverbals are functionally independent of one another, such that each may need to be established via direct reinforcement. The next step might therefore be to evaluate whether initial listener training facilitates subsequent intraverbal training by saving the total number of trials required to establish both repertoires, or vice versa. However, because prior research indicated that training of the reverse intraverbal relation might in fact produce listener relations (Luciano, 1986; Remington & Clark, 1993b; C. T. Sundberg & Sundberg, 1990), additional research on the effects of different types of intraverbal training may also be warranted. It also may be of both practical and conceptual interest to investigate under which circumstances intraverbal training generates an intraverbal relation in the opposite direction to that trained. Horne and Lowe (1996) have suggested that training an intraverbal relation frequently results in the establishment of the opposite intraverbal relation as well, due to self-echoic repetition during training. However, although this suggestion is central to some of the authors' analyses of emergent stimulus relations, they do not cite experimental data in its support. In the present study, reverse intraverbal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

categorization emerged following only one instance of intraverbal training, and similar findings have been obtained with typically developing adults (Polson & Parsons, 2000).

Additional areas for future research may be identified that relate to the sequencing of categorization skills in EIBI programs. For example, Maurice et al. (1996) recommend teaching category matching before teaching additional categorization skills. In the present study, category matching was not directly trained, and did not emerge as a result of either listener or intraverbal categorization training. It therefore might be worthwhile to evaluate the extent to which prior training in category matching facilitates subsequent training of other categorization skills. Based on the literature on stimulus equivalence, such results might be expected at least in the case of listener categorization.

In general, the literature on training intraverbal relations is relatively small. Researchers have pointed out the need for language interventions to focus on establishing intraverbal relations, as such relations appear to be a major component of typical verbal repertoires (e.g., Partington & Bailey, 1993; M. L. Sundberg & Michael, 2001). It has further been pointed out that research on intraverbals in conjunction with research on equivalence relations may be of applied as well as of theoretical importance (Stromer et al., 1996). Much of the research on intraverbal training to date (Braam & Poling, 1983; Luciano, 1986; Partington & Bailey, 1993; Miguel et al., 2005) has focused on teaching categorization skills, as did the present study. In EIBI curricula, however, intraverbal components are included in many programs other than those that teach categorization skills. Results obtained in the context of teaching categorization may certainly be applicable to teaching other types of intraverbal relations. However, it

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

may be noted that teaching intraverbal categorization necessarily involves training more than one intraverbal response to each verbal stimulus, or training multiple intraverbal responses to the same verbal stimulus. This is not the case with all types of intraverbal skills addressed in EIBI curricula. For example, some of the earliest programs that involve intraverbal training teach children to answer questions about animals and the sounds they make, such as /Which animal says meow?/ and /What does a cat say?/. In this type of training, only one intraverbal response is trained to each verbal stimulus. As noted earlier, this was the type of training employed in two of the studies in which intraverbal training generated listener relations (Remington & Clarke, 1993b; C. T. Sundberg & Sundberg, 1990). In those studies, the listener relations that emerged were analogous to an appropriate selection response to /Show me the one that says meow./ following the training of "cat" as a tact and as an intraverbal response to "Which animal says meow?". Future research might explore the extent to which it makes a difference whether the intraverbal relations trained are one-on-one relations, many-to-one relations, or one-to-many relations, and how such relations between stimuli and responses might interact with the direction of the intraverbal relation trained relative to the tested listener relation.

If among beginning language learners, functional independence of listener and intraverbal relations turns out to be a reliable finding, it might be worthwhile to investigate ways to overcome such functional independence. Since typically developing human adults appear capable of acquiring listener and intraverbal relations without direct reinforcement, it may be possible to bring the emergence of such relations under the control of whichever variables are responsible. In the present study, as noted before,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

it was not evaluated whether extended training or overtraining of one type of relation would ultimately bring about the emergence of the other, and this is one independent variable that might be investigated. Other avenues of research are possible as well. For example, some of the empirical literature on tacts and listener relations suggested that following a history of direct reinforcement of multiple instances of both tact and listener relations in the given training context, subsequent training of only one of the relations may suffice to establish both (e.g., Y. Barnes-Holmes et al., 2001a, 2001b; Guess & Baer, 1973; Smeets, 1978; Smeets & Striefel, 1976; see also Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005). Such multiple-exemplar instruction has produced similar results in other research involving topography-based verbal operants (e.g., Nuzzolo-Gomez & Greer, 2004). This type of instructional arrangement might be investigated in the context of intraverbal and listener relations as well.

Multiple-exemplar instruction, in fact, represents exactly the type of reinforcement history that has been hypothesized by proponents of relational frame theory (S. C. Hayes et al., 2001) to result in the ability to derive untrained speaker and listener relations (D. Barnes-Holmes et al., 2000). Other types of conceptual analyses suggest other types of reinforcement histories that might be relevant. Unpublished evidence appears to exist (Bell, 1999; cited in Horne et al., 2004) which indicates that the emergence of tacts as a result of listener training may be brought about by explicitly training the relation hypothesized by Horne and Lowe (1996) to mediate such emergence; that is, training children to emit an echoic response to a verbal stimulus as they select a nonverbal stimulus while looking at it. Future research might investigate

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the effects of directly training other performances that have been hypothesized to bring about emergence of speaker and listener relations (Horne & Lowe; Lowenkron, 1998).

On a final note, topography-based verbal relations have, as previously noted, rarely been the subject of basic research in the area of stimulus equivalence, even though such research has been called for (Hall & Chase, 1991; Polson & Parsons, 2000). As a result, not much in general is known about how the emergence of such relations is related to the patterns of emergent performances observed in this literature. This may be unfortunate given the widely held position that these phenomena are in some way related to human language (e.g., S. C. Hayes et al., 2001; Horne & Lowe, 1996). It appears that basic research is needed on the relationship among various topography-based verbal relations, as well as their relationship with listener relations. In addition to being of potential theoretical importance, such research would be likely to shed light on optimal strategies for teaching verbal relations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REFERENCES

Augustson, K. G., & Dougher, M. J. (1992). Teaching conditional discrimination to young children. *Experimental Analysis of Human Behavior Bulletin, 9*, 21-24.

Barnes, D., Browne, M., Smeets, P. M., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three- to six-year-old children. *The Psychological Record, 45,* 405-430.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's *Verbal Behavior*: A possible synthesis. *The Behavior Analyst, 23*, 69-84.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001a). Exemplar training and a derived transformation of function in accordance with symmetry. *The Psychological Record, 51*, 287-308.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001b). Exemplar training and a derived transformation of function in accordance with symmetry: II. *The Psychological Record, 51*, 589-603.

Boelens, H., Van den Broek, M., & Van Klarenbosch, T. (2000). Symmetric matching to sample in 2-year-old children. *The Psychological Record, 50*, 293-304.

Bones, R., Keenan, M., Askin, D., Adams, L., Taylor, D., & Nicholas, O. (2001). The effects of response topography on functional equivalence class formation. *The Psychological Record, 51*, 89-110.

Braam, S. J., & Poling, A. (1983). Development of intraverbal behavior in mentally retarded individuals through transfer of stimulus control procedures: Classification of verbal responses. *Applied Research in Mental Retardation, 4*, 279-302.

Brady, N. C., & McLean, L. K. (2000). Emergent symbolic relations in speakers and nonspeakers. *Research in Developmental Disabilities, 21*, 197-214.

Catania, A. C. (1998). *Learning* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Clarke, S., Remington, B., & Light, P. (1986). An evaluation of the relationship between receptive speech skills and expressive signing. *Journal of Applied Behavior Analysis, 19*, 231-239.

Clayton, M. C., & Hayes, L. J. (1999). Conceptual differences in the analysis of stimulus equivalence. *The Psychological Record, 49*, 145-161.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Connell, P. J. (1986). Acquisition of semantic role by language-disordered children: Differences between production and comprehension. *Journal of Speech and Hearing Research, 29*, 366-374.

Connell, P. J., & McReynolds, L. (1981). An experimental analysis of children's generalization during lexical learning: Comprehension or production. *Applied Psycholinguistics, 2*, 309-332.

Cuvo, A. J., & Riva, M. T. (1980). Generalization and transfer between comprehension and production: A comparison of retarded and nonretarded persons. *Journal of Applied Behavior Analysis, 13*, 315-331.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence-class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46*, 243-257.

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 59*, 103-114.

Eikeseth, S., & Jahr, E. (2001). The UCLA reading and writing program: An evaluation of the beginning stages. *Research in Developmental Disabilities, 22*, 289-307.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 3*, 345-358.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, 2, 121-135

Galvao, O. D., Calcagno, S., & Sidman, M. (1992). Testing for emergent performances in extinction. *Experimental Analysis of Human Behavior Bulletin, 10*, 18-20.

Goldstein, H., Angelo, D., & Mousetis, L. (1987). Acquisition and extension of syntactic repertoires by severely mentally retarded youth. *Research in Developmental Disabilities, 8*, 549-574.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229-262). New York: Plenum.

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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Guess, D. (1969). A functional analysis of receptive language and productive speech: Acquisition of the plural morpheme. *Journal of Applied Behavior Analysis, 2*, 55-64.

Guess, D., & Baer, D. M. (1973). An analysis of individual differences in generalization between receptive and productive language in retarded children. *Journal of Applied Behavior Analysis, 6*, 311-329.

Hall, G., & Chase, P. N. (1991). The relationship between stimulus equivalence and verbal behavior. *The Analysis of Verbal Behavior, 9*, 107-119.

Hayes, L. J., Tilley, K. J., & Hayes, S. C. (1988). Extending equivalence class membership to gustatory stimuli. *The Psychological Record, 38*, 473-482.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition.* New York: Plenum.

Holdgrafer, G. (1981). Mode-relations in language learning by language-deficient retarded subjects. *Perceptual and Motor Skills, 53*, 520-522.

Holdgrafer, G., & McReynolds, L. (1975). An experimental analysis of comprehension and production in children's acquisition of morphological rules. *Human Communication, 5*, 45-62.

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., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young children: IV: Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior, 85*, 247-273.

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.

Jordan, C. R., Pilgrim, C., & Galizio, M. (2001). Conditional discrimination and stimulus equivalence in young children following three different baseline training procedures. *Experimental Analysis of Human Behavior Bulletin, 19*, 3-7.

Keller, M. F., & Bucher, B. D. (1979). Transfer between receptive and productive language in developmentally disabled children [Abstract]. *Journal of Applied Behavior Analysis, 12*, 311.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Kennedy, C. H., Itkonen, T., & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis, 27*, 673-683.

Leaf, R., & McEachin, J. (1999). *A work in progress. Behavior management strategies and a curriculum for intensive behavioral treatment of autism.* New York: DRL.

Lee, V. L. (1981). Prepositional phrases spoken and heard. *Journal of the Experimental Analysis of Behavior, 35*, 227-242.

Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology, 56*, 201-239.

Lovaas, O. I. (1977). *The autistic child.* New York: Irvine.

Lovaas, O. I. (1981). *Teaching developmentally disabled children. The ME book.* Austin, TX: PRO-ED.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3-9.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays. Basic intervention techniques.* Austin, TX: PRO-ED.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior, 78*, 527-549.

Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior, 83*, 47-65.

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior, 69*, 327-354.

Luciano, M. C. (1986). Acquisition, maintenance, and generalization of productive intraverbal behavior through a transfer of stimulus control procedures. *Applied Research in Mental Retardation, 7*, 1-20.

Mackay, H. A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis and Intervention in Developmental Disabilities, 5*, 373-387.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Mackay, H. A., & Ratti, C. A. (1990). Position-numeral equivalences and delayed position recognition span. *American Journal on Mental Retardation, 95*, 271-282.

Maurice, C., Green, G., & Luce. S. (Eds.). (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals.* Austin, TX: PRO-ED.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. *The Analysis of Verbal Behavior, 3*, 1-4.

Michael, J. (1996). Separate repertoires or naming? *Journal of the Experimental Analysis of Behavior, 65*, 296-298.

Miguel, C. M., Petursdottir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior, 21*, 27-41.

Miller, M. A., Cuvo, A. J., & Borakove, L. (1977). Teaching naming of coin values – Comprehension before production versus production alone [Abstract]. *Journal of Applied Behavior Analysis, 10*, 735-736.

Nuzzolo-Gomez, R., & Greer, R. D. (2004). Emergence of untaught mands or tacts of novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior, 20*, 63-76.

O'Donnell, J., & Saunders, K. J. (2003). Equivalence relations in individuals with language limitations and mental retardation. *Journal of the Experimental Analysis of Behavior, 80*, 131-157.

Partington, J. W., & Bailey, J. S. (1993). Teaching intraverbal behavior to preschool children. *The Analysis of Verbal Behavior, 11*, 9-18.

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior, 73,* 177-193.

Polson, D. A. D., & Parsons, J. A. (2000). Selection-based versus topography-based responding: An important distinction for stimulus equivalence? *The Analysis of Verbal Behavior, 17*, 105-128.

Potter, B., & Brown, D. L. (1997). A review of studies examining the nature of selection-based and topography-based verbal behavior. *The Analysis of Verbal Behavior, 14*, 85-104.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remington, B., & Clarke, S. (1993a). Simultaneous communication and speech comprehension: I. Comparison of two methods of teaching expressive signing and speech comprehension skills. *Augmentative & Alternative Communication, 9*, 36-48.

Remington, B., & Clarke, S. (1993b). Simultaneous communication and speech comprehension: Part II. Comparison of two methods of overcoming selective attention during expressive sign training. *Augmentative & Alternative Communication, 9*, 49-60.

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency, 77,* 515-523.

Sidman, M., Cresson, O., & Wilson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior, 22*, 261-273.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5-22.

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.

Skinner, B.F. (1957). *Verbal behavior.* Acton, MA: Copley.

Smeets, P. M. (1978). Establishing generative performance and cross modal generalization of the manual plural sign in a severely retarded deaf girl. *British Journal of Communication Disorders, 13*, 49-57.

Smeets, P. M., & Barnes-Holmes, D. (2005). Auditory-visual and visual-visual equivalence relations in children. *The Psychological Record, 55*, 483-503.

Smeets, P. M., & Striefel, S. (1976). Acquisition and cross modal generalization of receptive and expressive signing skills in a retarded deaf girl. *Journal of Mental Deficiency Research, 20*, 251-259.

Stromer, R., Mackay, H. A., & Remington, B. (1996). Naming, the formation of stimulus classes, and applied behavior analysis. *Journal of Applied Behavior Analysis, 29*, 409-431.

Sundberg, C. T., & Sundberg, M. L. (1990). Comparing topography-based verbal behavior with stimulus selection-based verbal behavior. *The Analysis of Verbal Behavior, 8*, 31-41.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Sundberg, M. L., & Michael, J. (2001). The benefit's of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25*, 698-724.

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities.* Pleasant Hill, CA: Behavior Analysts, Inc.

Watters, R. G., Wheeler, L. J., & Watters, W. E. (1981). The relative efficiency of two orders for training autistic children in the expressive and receptive use of manual signs. *Journal of Communication Disorders, 14*, 273-285.

Williams, G. C., & McReynolds, L. V. (1975). The relationship between discrimination and articulation training in children with misarticulations. *Journal of Speech and Hearing Research, 18*, 401-412.

Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions, 18*, 245-266.

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior, 78*, 237-248.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix A

HSIRB Approval Letter

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

WESTERN MICHIGAN UNIVERSITY

Human Subjects Institutional Review Board

Date: January 3, 2005

To: James Carr, Principal Investigator
Anna Petursdottir, Student Investigator for dissertation
Sarah Lechago, Student Investigator

From: Amy Naugle, Ph.D., Interim Chair Amy Naugle

Re: HSIRB Project Number: 04-12-03

This letter will serve as confirmation that your research project entitled "An Evaluation of Intraverbal Training, Listener Training and Category Matching for Teaching Categorization Skills" has been **approved** under the **full** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may **only** conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: December 15, 2005

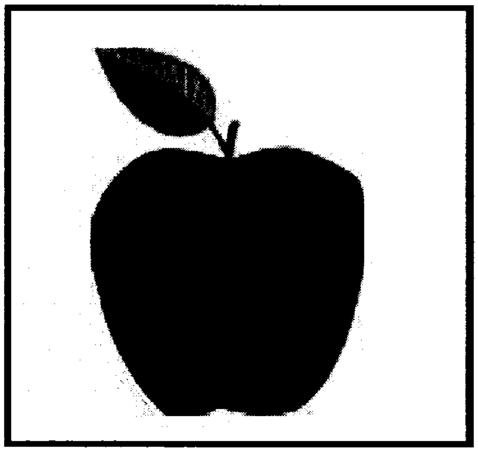
Walwood Hall, Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276

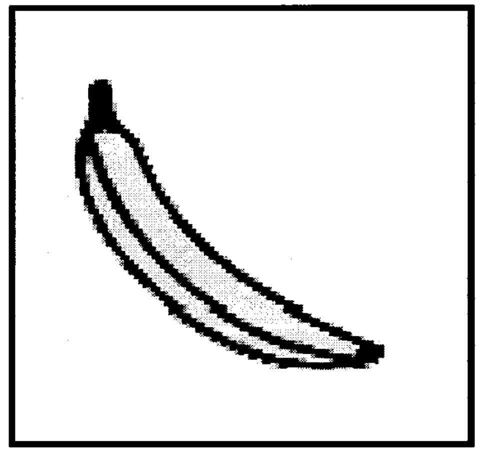
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

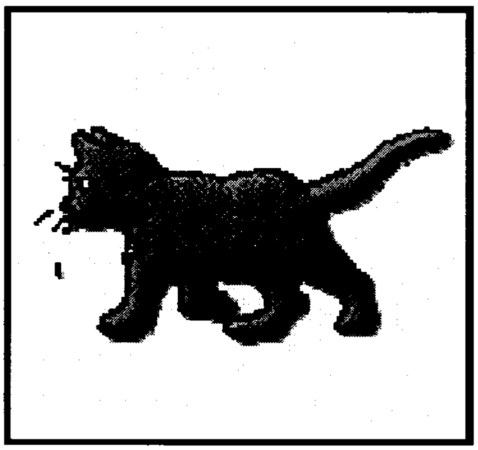
Appendix B

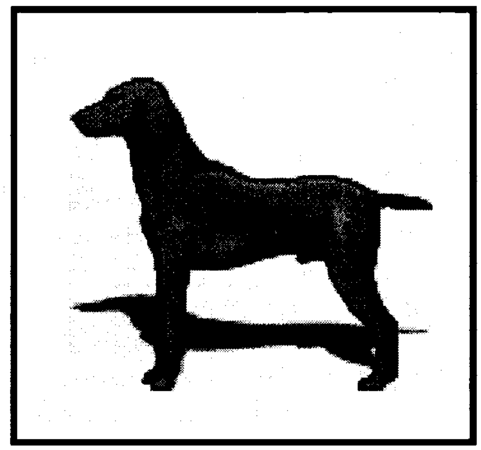
Pretraining Stimuli

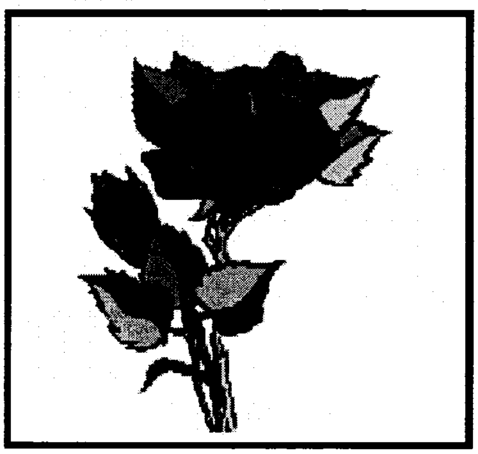
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

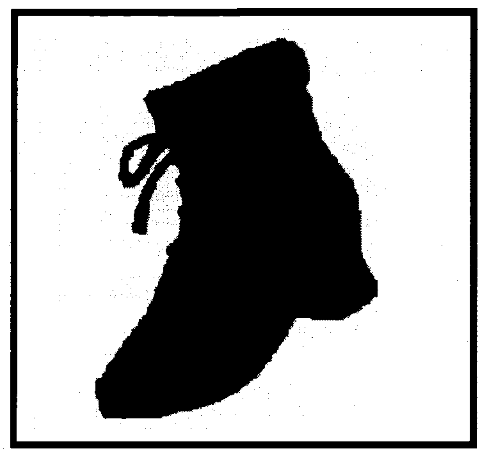










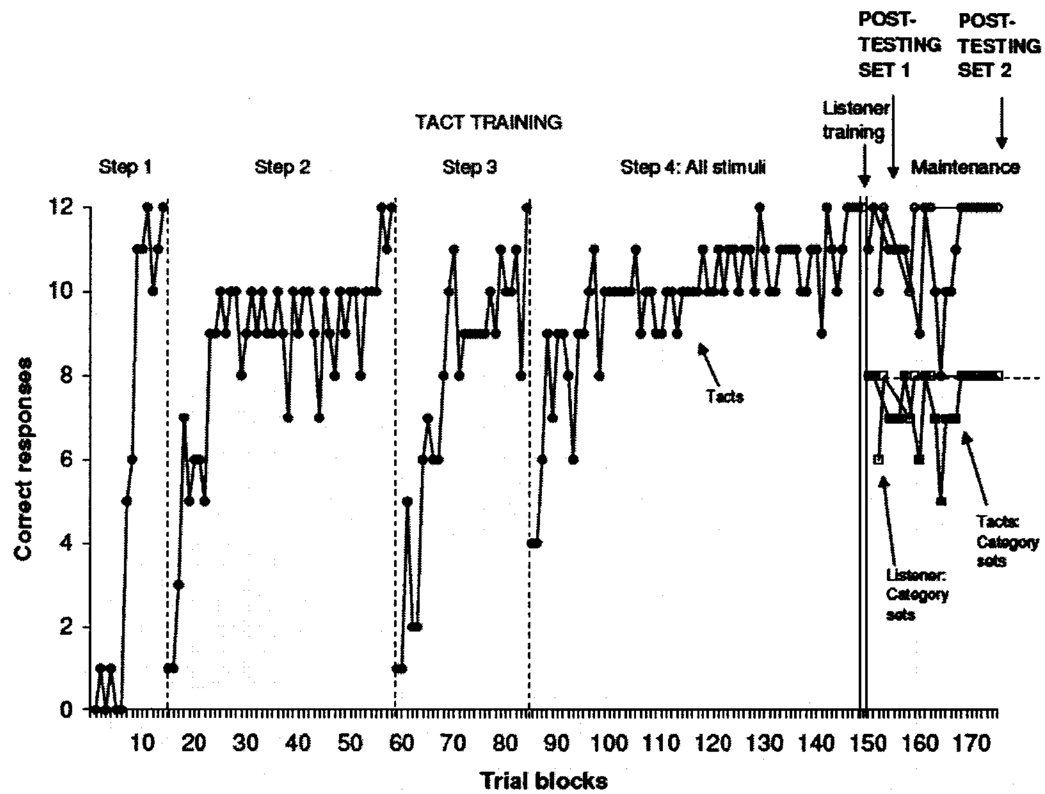


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix C

Exemplar Name Tact and Listener Training and Maintenance Data

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Tact and listener training and maintenance data for Marc on the East/West and Greek/Cyrillic category sets, and additional stimuli "Al", "Mem", "Chile", and "Tina". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

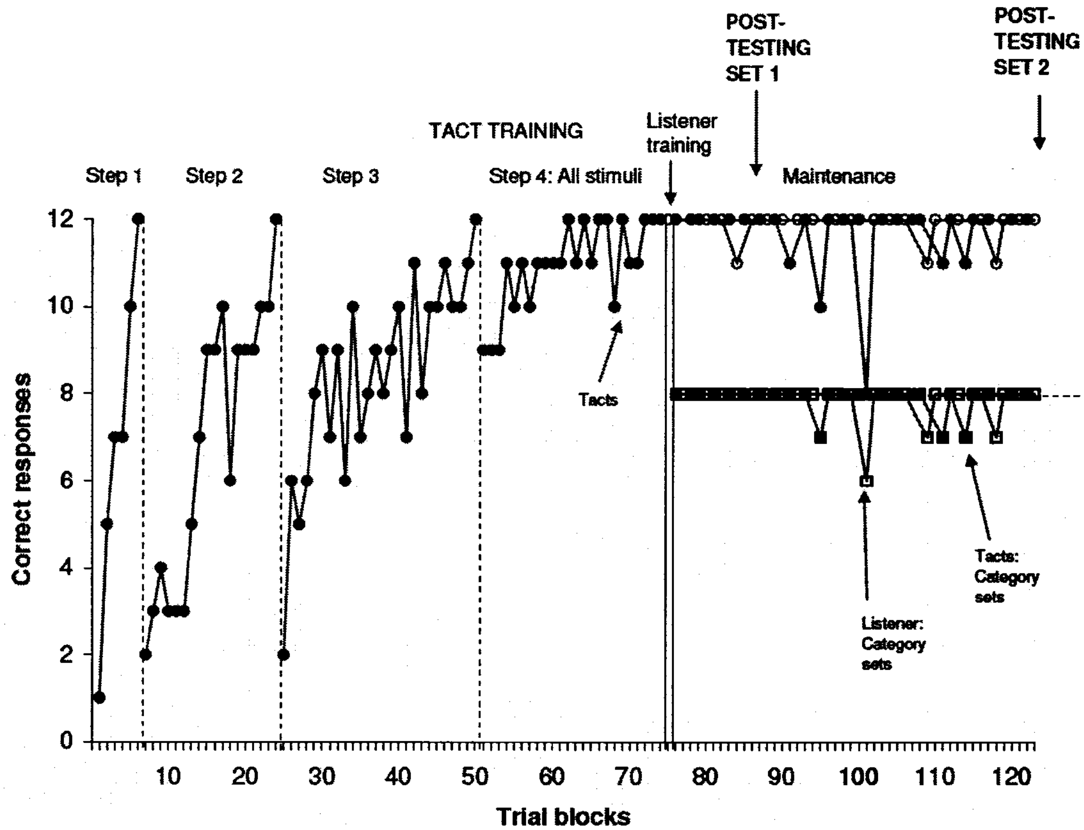
<u>Training</u>

- Agreement assessed on 82.6% of all blocks
- Average agreement: 99.8%
- Range: 83.3-100.0%

<u>Maintenance</u>

- Agreement assessed on 33.3% of all blocks
- Average agreement: 100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Tact and listener training and maintenance data for Greg on the North/South and Kata/Hira category sets, and additional stimuli "Hey", "Kaf", "Fiji" and "Nepal". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

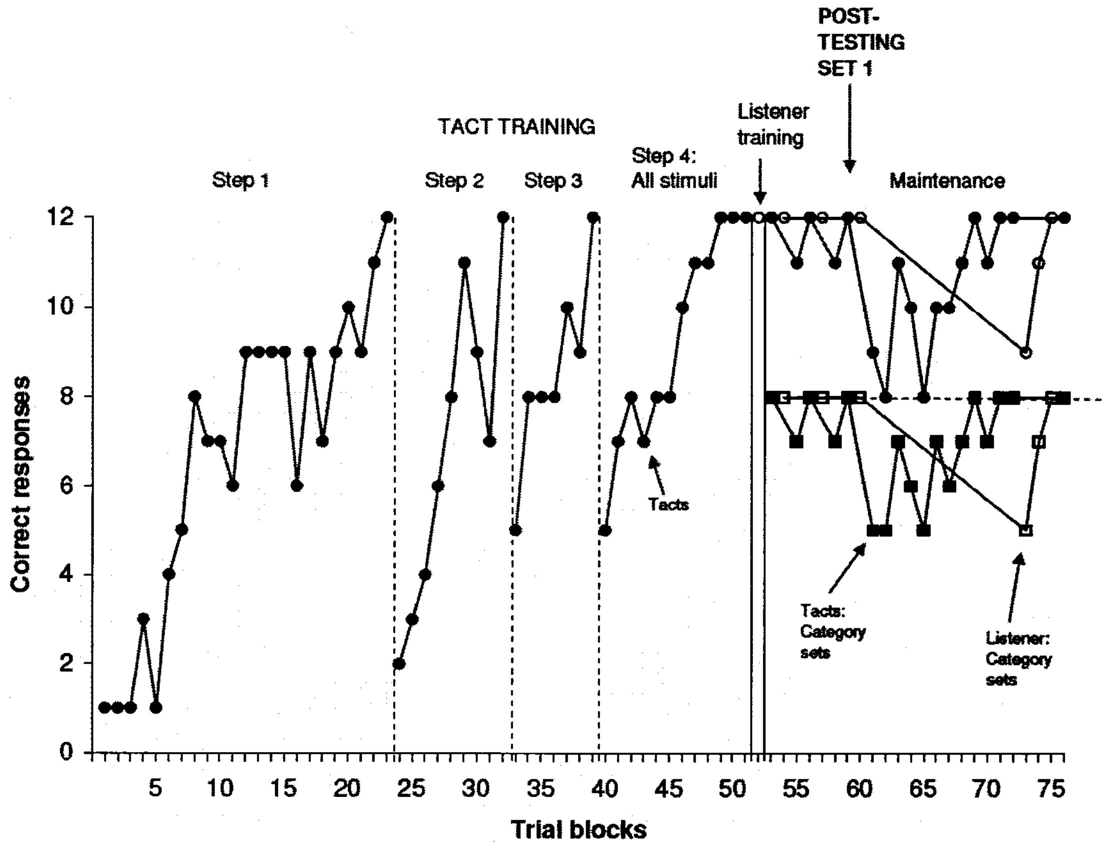
<u>Training</u>

- Agreement assessed on 58.6% of all blocks
- Average agreement: 99.6%
- Range: 91.7-100.0%

<u>Maintenance</u>

- Agreement assessed on 36.7% of all blocks
- Average agreement: 99.5%
- Range: 91.7-100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Tact and listener training and maintenance data for Sam on the East/West and Greek/Cyrillic category sets, and additional stimuli "Al", "Mem", "Chile", and "Tina". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

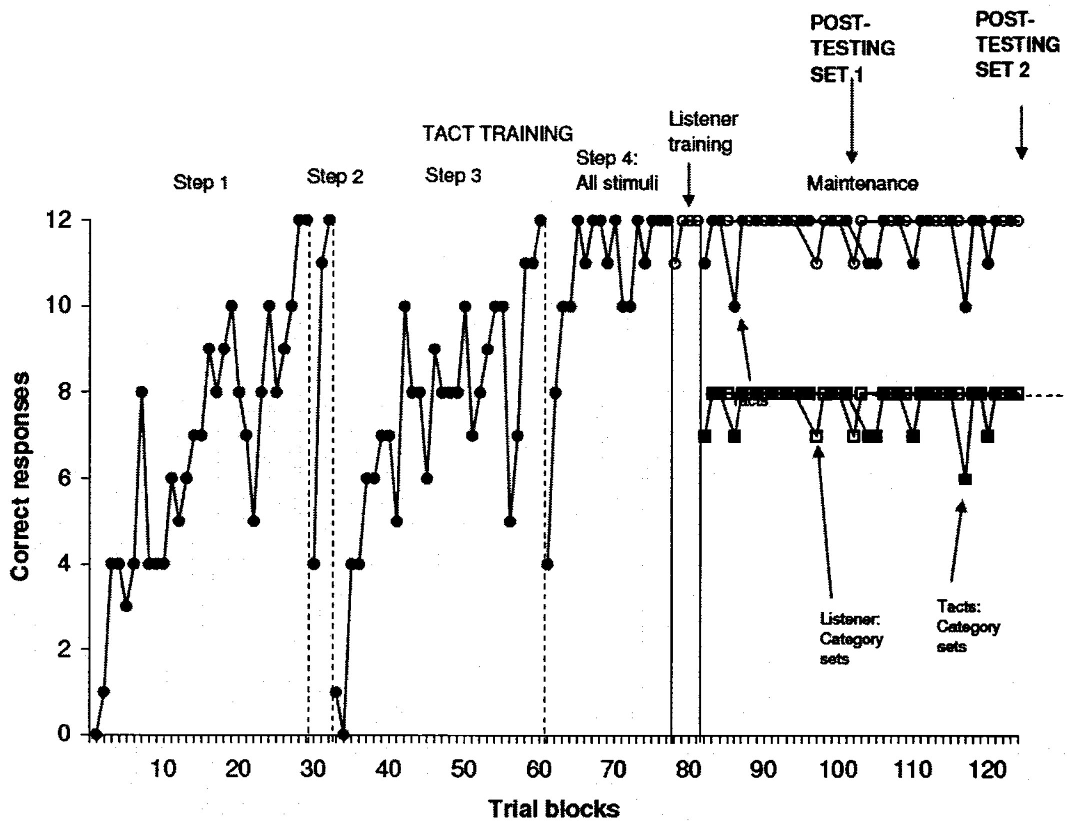
Training

- Agreement assessed on 29.3% of all blocks
- Average agreement: 100.0%

Maintenance

- Agreement assessed on 29.2% of all blocks
- Average agreement: 98.8%
- Range: 91.7-100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Tact and listener training and maintenance data for Erika on the East/West and Greek/Cyrillic category sets, and additional stimuli "Al", "Mem", "Chile", and "Tina". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

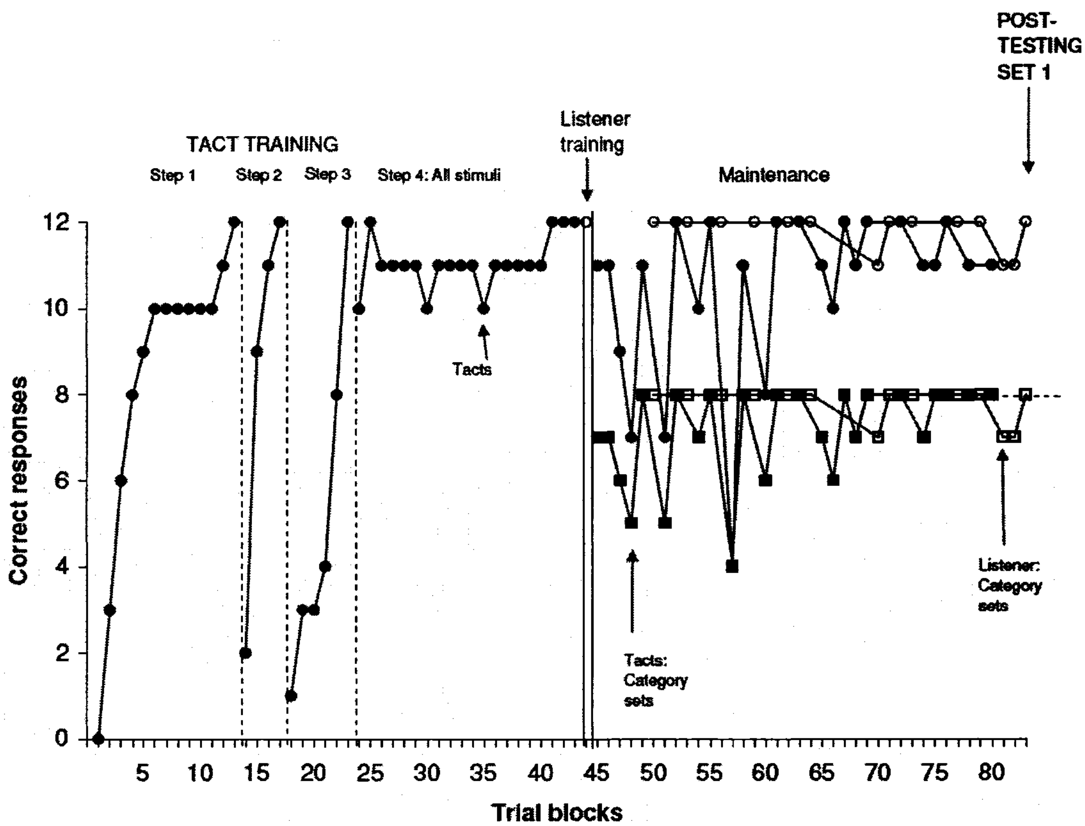
<u>Training</u>

- Agreement assessed on 75.3% of all blocks
- Average agreement: 99.9%
- Range: 91.7-100.0

<u>Maintenance</u>

- Agreement assessed on 47.7% of all blocks
- Average agreement: 100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



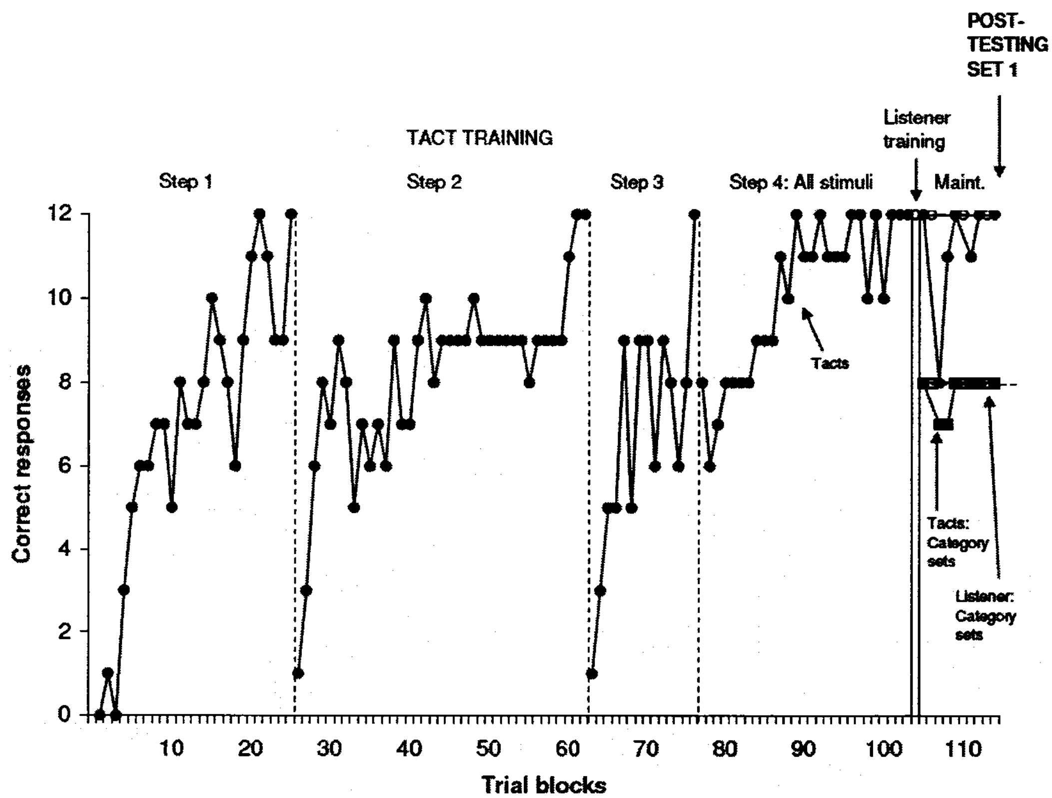
Tact and listener training and maintenance data for Chanele on the North/South and Kata/Hira category sets, and additional stimuli "Hey", "Kaf", "Fiji" and "Nepal". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

Training
- Agreement assessed on 34.1% of all blocks
- Average agreement: 100.0%

Maintenance
- Agreement assessed on 62.7% of all blocks
- Average agreement: 100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Tact and listener training and maintenance data for Sanjay on the North/South and Kata/Hira category sets, and additional stimuli "Hey", "Kaf", "Fiji" and "Nepal". The maintenance phase depicts performance on all listener trials (open circles), listener trials for stimuli belonging to categorization training sets only (open squares), all tact trials (filled circles) and tact trials for stimuli belonging to categorization training sets only (filled squares).

Point-by-point interobserver agreement data across trial blocks.

<u>Training</u>

- Agreement assessed on 81.7% of all blocks
- Average agreement: 99.9%
- Range: 91.7-100.0%

<u>Maintenance</u>

- Agreement assessed on 60.0% of all blocks
- Average agreement: 100.0%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.