Evaluating the Effect of Physical Prompts on Echoic Responses

Kristin S. Bowman

Endicott College

Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Department of Applied Behavior Analysis

Endicott College, Beverly, MA

May 2023

Abstract

Kristin S. Bowman

Dissertation Advisor: Mary Jane Weiss

Failure to acquire vocal communication is one of the most pressing concerns for parents of children on the autism spectrum. While vocal speech is often preferred to alternate communication modalities, and prioritized in treatment, its social significance is further supported by its role as a strong predictor of later verbal abilities and association with greater independence in adulthood. Approximately 1 in 4 children on the autism spectrum present with limited to absent vocal–verbal communication, even after intervention. Speech sound production and vocal imitation remain relatively understudied areas in autism treatment, yet if we are to increase the number of individuals on the autism spectrum who are proficient vocal communicators, we must develop impactful, effective treatments that enable spoken language. In the current study, conventional vocal imitation training was compared to the use of physical prompts for increasing echoic responses in children on the autism spectrum with speech sound disorders. All participants showed improvements following intervention, but the children with more restricted echoic repertoires only acquired the target vocalizations that were taught using physical prompts. These findings suggest that physical prompts may be a viable treatment for improving echoic behavior in children on the autism spectrum, especially those with minimal to absent vocal communication.

Keywords: echoics, physical prompts, autism, vocal imitation training, vocal communication

Table of Contents

Committee and Acknowledgements

Committee Chair: Mary Jane Weiss

Committee Members: Joanne Gerenser, Lina Slim, and Kathleen Dyer

I am honored to have a committee of female leaders who have devoted much of their career to improving interprofessional collaboration, particularly between speech-language pathology and applied behavior analysis. Thank you to my committee members, Drs. Joanne Gerenser, Lina Slim, and Kathleen Dyer, for your encouragement and guidance on this research. As you have demonstrated in your own work, combining the sciences of both human behavior and communication often bears a product greater than either discipline could accomplish independently. It is my hope that this project further highlights opportunities for collaborative efforts and benefits both fields of speech-language pathology and behavior analysis in research and practice.

I owe a special thank you to Dr. Mary Jane Weiss, my mentor and advisor. You have an uncanny way of making the most arduous tasks seem effortless. Through your leadership I have learned that every problem has a solution, that a perfect blend of patience and tenacity are often all that is needed for success, and that superheroes are real! Thank you for your endless optimism and for the opportunities you created for me.

My sincere thanks to all those who helped with interobserver agreement, procedural fidelity, and editing the manuscript: Dr. Rebekah Lee, Dr. Natalie Driscoll, Courtney Keleher, and Nicole Radzilowicz. Thank you to my very own, Cohort Five. I am so proud to know each and every one of you! Your brilliance, love, support, and never-ending humor made this an unforgettable journey. I would like to especially thank Drs. Kimberly Marshall, Lisa Tereshko, and Videsha Marya for always listening and offering their expert feedback.

Finally, I am forever grateful to my loving family who made many sacrifices along the way. To my mom and dad, you have always cheered me on and helped with anything I ever needed, even at a moment’s notice. To my sweet children: Sam, Sadie, Nick, and Rosie. Thank you for the many ways you supported me on this adventure. I will never forget how you took on extra chores, so I had time to study, and all the coffee refills and meals you served me at the computer, and even the time you spent checking my citations! I love you with all my heart. And to my husband Brian, the love of my life. Thank you for helping make my dreams come true. Thank you for reminding me that I can do it! You always believe in me. It is truly a blessing to share my life with you.

List of Figures

List of Tables

Evaluating the Effect of Physical Prompts on Echoic Responses

Autism spectrum disorder is one of the most prevalent developmental disabilities (Zablotsky et al., 2019) with approximately 1 in 100 children diagnosed worldwide (Zeidan et al., 2022) and as many as 1 in 36 diagnosed in the United States (Maenner et al., 2023). Persistent deficits in social communication are one of its defining characteristics, (American Psychiatric Association, 2013) yet the communication profiles of individuals on the autism spectrum are quite diverse (Norrelgen et al., 2015). Although several studies have indicated that many individuals on the autism spectrum communicate vocally in 2+ word phrases, a fair number fail to acquire functional vocal verbal communication. For example, in their longitudinal assessment of verbal skills in 98 children on the autism spectrum, Anderson et al. (2007) estimated expressive language abilities according to the specific module of the *Autism Diagnostic Observation Schedule* (ADOS) selected for administration at 9 years of age. Nearly 24% of the children were verbally fluent and spoke in complex sentences, 23.8% spoke in phrases but were not verbally fluent, and another 23.8% communicated with words, but not phrases. A slightly larger percentage, 28.6%, had no vocal words or very few inconsistent words at 9 years of age. As another example, Norrelgen et al. (2015) evaluated 97 children, ages 4 to 6.5 years, to identify the distribution of verbal abilities in children on the autism spectrum. After two years of early intensive behavioral services, 58.8% of the children used 2–word phrases and had an expressive age equivalency of 24+ months, 16.5% used at least three words but rarely spoke in 2–word phrases and had an expressive age equivalency of <24 months, and 24.7% used fewer than three words and had an expressive age equivalency of <15 months. Similarly, Rose et al. (2016) sought to identify the proportion of minimally verbal children on the autism spectrum between 30 and 71 months of age. Using a comprehensive battery of autism diagnostic tools, direct standardized language assessments, and parent–report of communication, they evaluated 246 children on the autism spectrum. After an average of 14 months of early intensive behavioral intervention, 26.3% of the children had less than five spontaneous and functional words. Combined, results from these studies indicate that even after intervention, approximately 1 in 4 children on the autism spectrum will have limited to absent vocal verbal communication.

Vocal Communication as a Prognostic Indicator and/or Predictive Factor

Several longitudinal studies have examined early childhood developmental profiles to identify prognostic indicators and/or predictive factors associated with autism spectrum disorder and the acquisition of spoken language (Siller & Sigman, 2008). According to Adolfsson and Steineck (2000), both prognostic indicators and predictive factors are used to foresee outcomes associated with a particular condition and typically include specific patient characteristics (e.g., age, gender) and/or aspects of the disorder (e.g., severity, developmental skills). However, prognostic indicators refer more specifically to variables that predict outcomes of the condition alone, while predictive factors are those that foresee the effects of treatment on the condition (Adolfsson & Steineck, 2000). Although various factors that may influence spoken language in autism have been investigated, including age, severity of symptoms, comorbid conditions, developmental regression, joint attention, nonverbal cognition, play, parent responsiveness, and social motivation, (e.g., Bopp et al., 2009; Norrelgen et al., 2015; Pecukonis et al., 2019; Sigman & Ruskin, 1999; Siller & Sigman, 2008; Su et al., 2021; Thurm et al., 2007; Yoder et al., 2015) the results are inconsistent. Autism is a complex neurodevelopmental disorder, and there are likely many contributing variables (Beiting, 2022; Pecukonis et al., 2019). Nonetheless, vocal communication is of great social significance. If we are to increase the number of individuals on the autism spectrum who are proficient vocal communicators, we must identify predictive factors and develop impactful, effective treatments that enable spoken language, especially for those with minimal to absent vocal communication (Chenausky et al., 2018; Saul & Norbury, 2020; Yoder et al., 2015).

Many of these longitudinal investigations have examined the influence of speech production on later expressive language or overall verbal abilities. For example, Wetherby et al. (2007) compared the social–communication profiles of 125 children, 18–24 months of age: 25 with developmental delays, 50 typically developing (TD), and 50 later diagnosed with autism spectrum (LDA) at 3 years of age. These social communication profiles were further examined to determine what variables, if any, predicted verbal abilities in children on the autism spectrum at 3 years of age. While consonant production for the LDA children and those with developmental delays was comparable, significant differences were displayed by the LDA children and the TD children. Additionally, a large positive correlation between consonant inventory and number of words (i.e., words or approximated words spoken or signed) was observed. The strongest predictors of verbal abilities at 3 years of age, based on an average of scores from the expressive language and receptive language subtests of the *Mullen Scales of Early Learning,* were understanding, acts for behavior regulation, and inventory of consonants at 18–24 months of age.

Using the same cohort of participants as Wetherby et al. (2007), Plumb and Wetherby (2013) further compared the vocal behaviors of the children in each group. The *Communication and Symbolic Behavior Scales Developmental Profile* was administered to all children at 18–24 months of age. This assessment was video recorded, and the children's vocalizations were later coded for precise measures of vocal behavior. The LDA children emitted a lower proportion of transcribable vocalizations (i.e., syllable marked by at least one vowel and may or may not contain a consonant) and a higher proportion of nontranscribable vocalizations (i.e., no recognizable vowel or vowel with atypical phonation), as compared to the other groups. Additionally, the authors found a positive correlation between the proportion of transcribable vocalizations and scores on the receptive and expressive scales from the *Mullen Scales of Early Learning* administered at 3 years of age. That is, vocalizations marked by the presence of true vowels, consonants, and syllables in the second year of life were associated with higher language performance at 3 years of age.

Saul and Norbury (2020) found phonetic repertoire and consonant inventory to be the strongest predictors of expressive language abilities (i.e., number of words spoken). The study included 27 children with a diagnosis of autism spectrum disorder. All children were 2–5 years of age and described as “minimally verbal,” which was defined as an expressive vocabulary of less than 24 spoken words. Expressive language was measured at the start of the study and again 12 months later. Changes in expressive language were highly variable although most children still presented as minimally verbal 12 months later. Phonetic repertoire (i.e., reported, observed, and elicited speech sounds) and consonant inventory (i.e., speech sounds produced as part of intentional communicative acts) were stronger predictors of expressive language than alphabet and phonics score, and socio–communicative measures such as intentional communication, response to joint attention, and parent responsiveness.

Paul et al. (2008) evaluated 37 children on the autism spectrum at 15–25 months of age and, at least one year later, when the children were 36–58 months of age. A comprehensive battery of assessments, including direct observation and indirect measures of development, were administered at both time points. Significant deficits in expressive language, as compared to nonverbal abilities, were noted. To further examine gains in expressive language over time, the participants were divided into two groups based on scores on the *Vineland Adaptive Behavior Expressive Communication Scale* at the follow–up assessment: (1) those with relatively better language outcomes (i.e., age equivalency above 30 months of age), and (2) those with relatively worse language outcomes (i.e., age equivalency below 30 months of age). Results of the initial assessments for each group were further analyzed. The children with higher expressive language abilities at the follow–up assessment presented with higher nonverbal cognition, receptive language, use of play schemes, response to joint attention, and vocal production of sounds and words at the initial assessment in the second year of life. In other words, the children with poorer expressive language outcomes at 36–58 months of age showed lower performance on receptive language tasks and emitted fewer speech–like sounds or word approximations at 15–25 months of age.

Chenausky et al. (2018) compared predictors for speech improvement such as gender, chronological age, expressive language, phonetic inventory, autism severity, and nonverbal intelligence. The study included 38 children on the autism spectrum, between ages 3 years, 5 months and 10 years, 8 months, who spoke fewer than 20 intelligible words. They found phonetic inventory at baseline (i.e., the number of English phonemes correctly imitated) to be the strongest predictor of speech improvement. Because nonverbal intelligence, expressive language, and age did not reliably predict gains in speech production following treatment, the authors recommended that older children with autism be considered viable candidates for speech therapy.

Biller and Johnson (2019) examined the social–cognitive abilities and speech sound production in five children on the autism spectrum, ages 3 years, 7 months to 6 years, 11 months. All the children were described as minimally verbal (i.e., spontaneous expressive vocabulary of less than 25 words) and relied on alternative and augmentative means of communication such as sign language, picture exchanges, or a speech generating device. Several areas of social–cognition were measured including communicative intent, joint attention, symbolic play, and gestural use via the *Communication and Symbolic Behavior Scales*. The social–cognition abilities of all children fell within the 12–24 month age range. The Focal Oromotor Control section of the *Verbal Motor Production Assessment for Children* was administered to evaluate the children's speech sound production. Specifically, measures included onset of voicing, imitation of vowels and consonants in isolation, vowel sequences, and various consonant–vowel combinations. Performance on the speech sound production assessment yielded correct responding between 4% and 54% of tasks. According to the *Framework for Spoken Language*, all the children functioned at the prelinguistic or first words stage of communication development. Deficits in spoken language have generally been attributed to social–cognitive abilities, but the authors noted that the children's scores on the assessments of social–cognition did not correspond to their number of words spoken. Results were in support of Wetherby et al.'s (2007) study, indicating that speech sound production was more closely associated with the number of words spoken.

Yoder et al. (2015) evaluated 87 children on the autism spectrum over a 16–month period. Participants were 24–48 months of age at the start of the study and were described as minimally verbal in that they reportedly produced no more than 20 words in total and were observed to emit no more than 5 root words during a 15–min observation. At the final assessment, 16 months later, 40% of the participants were still classified as minimally verbal. The authors found joint attention, intentional communication, parent linguistic responses, and consonant inventory to be the strongest predictors of expressive language growth.

Thurm et al. (2007) evaluated the expressive and receptive language of 59 children on the autism spectrum at 2 years of age and again at 5 years of age. Comparing various predictive factors, they found that both nonverbal cognitive ability and language at 2 years of age reliably predicted expressive language at 5 years of age. To identify additional factors associated with significantly limited language at 5 years of age, the authors selected the children who had little to no expressive language at 2 years of age and a nonverbal cognitive age equivalency of 18+ months. It was expected that a child with this nonverbal ability would be capable of emitting at least five spontaneous words daily. This group of children was further divided into two groups: (1) some language at 5 years of age, and (2) no language at 5 years of age (as measured by the overall language level of the *Autism Diagnostic Interview–Revised*). Performance on standardized assessments at 2 years of age was compared. Significant differences between the two groups were observed across multiple variables including nonverbal cognitive ability, socialization, expressive and receptive communication, responding to joint attention, initiating joint attention, imitating speech sounds, and imitating simple motor actions following adult model. That is to say, children with relatively higher nonverbal cognition, but who failed to develop language at 5 years age, presented with more impaired joint attention, motor imitation, and vocal imitation at 2 years of age.

Smith et al. (2007) evaluated 35 children on the autism spectrum four times over a period of 2 years to examine vocabulary growth and the predictive value of chronological age, developmental level, autism severity, and early language skills. The children had a mean age of 45 months (range 20–71 months) at the start of the study and an expressive vocabulary of fewer than 60 words. Developmental assessment tools included the *MacArthur–Bates Communicative Development Inventory*, *Childhood Autism Rating Scales*, and the *Mullen Scales of Early*

Learning. Children were divided into four clusters based on the number of words produced according to the *MacArthur–Bates Communicative Development Inventory.* Cluster 1 was comprised of 15 children with the slowest vocabulary growth. Their vocabulary increased an average of only 9.72 words (range 0–56 words) following 24 months of intervention. Cluster 2 included eight children whose vocabulary increased an average of 200.25 words (range 139–314 words) after 24 months of intervention. Cluster 3 included seven children who demonstrated an average growth of 453.43 words (range 399–625 words) and Cluster 4 consisted of only five children, each with steep rates of growth, acquiring an average of 638 words (range 646–697 words) over the 2–year period. Examining differences among the groups at the initial assessment, the authors found no significant differences in chronological age, autism severity, scores on the Early Learning Composite from the *Mullen Scales of Early Learning,* phrases understood, words comprehended, or use of symbolic gestures. Instead, the strongest predictors of rapid expressive vocabulary growth after 2 years of intervention were spoken vocabulary, pretend play with objects, the number of gestures to initiate joint attention, and verbal imitation skills.

Given the heterogeneity in spoken language within the autism population, it is unlikely that one single factor can explain the observed impairments in vocal communication (Tager–Flushberg & Kasari, 2013). However, these studies highlight the important relationship between speech production and language development in children on the autism spectrum. Learning to vocalize is likely met with feedback from caregivers and communication partners, thus increasing social opportunities for language acquisition and improving later verbal abilities (Plumb & Wetherby, 2013). Useful speech before 5 years age (i.e., composite language score of 0–2 on the *Autism Diagnostic Interview*) has also been associated with improved outcomes in adulthood in areas such as verbal ability, adaptive skills, academic achievement, residential status, and overall independence (Howlin et al., 2004; Venter et al., 1992). Certainly, vocal communication offers the most efficient means of exchange (Stone & Yoder, 2001). Failure to acquire spoken language is one of the most pressing concerns reported by parents of children on the autism spectrum (Mayo et al., 2013) and is preferred over alternate modalities of communication (see Butt et al., 2022; Romski & Sevcik, 2005). Speech sound production is indeed necessary for vocal verbal communication (Biller & Johnson, 2019), but the exact reasons many children on the autism spectrum fail to acquire functional vocal speech is unknown (Beiting, 2022; Rogers et al., 2006; Tager–Flushberg & Kasari, 2013).

Acquisition of Spoken Language

Vocal development begins in infancy (Oller et al., 1999). Even before the emergence of true words, infants engage in vocal behaviors, which systematically change during the first year of life and serve as important precursors to more mature speech patterns. The development of vocal behaviors during this prelinguistic period are dependent on both anatomical changes and exposure to the specific sounds and prosodic features of the native language (Kuhl & Meltzoff, 1996). As the vocal tract and articulators develop, children learn to imitate the sounds they hear. This prelinguistic period ends with the appearance of the first word, and the linguistic period follows where phonological, lexical, and syntactic development continue (Bauman–Waengler, 2004; Hoff, 2001).

Prelinguistic Period

Prelinguistic vocal development is often classified into five, overlapping stages which will be further described (Hoff, 2001; Nathani et al., 2006; Ramsdell–Huddock et al., 2019, Stark, 1980).

Reflexive phonation occurs from birth to 2 months of age (Nathani et al., 2006; Patten et al., 2014; Ramsdell–Huddock et al., 2019; Stark, 1980). The vocal behaviors here are characterized by crying, grunting, low pitch quasi–vowels, and vegetative sounds such as burping, coughing, grunting, and sneezing. At this stage, vocalizations are reflexive, occurring in response to activity and biological processes such as feeding (Ramsdell–Huddock et al., 2019).

Between 1 and 4 months of age, infants enter the *primary articulation* stage (Hoff, 2001; Ramsdell–Huddock et al., 2019) where vocalizations are no longer simply reflexive responses to physiological needs, but the infant learns to control phonation (Nathani et al., 2006). During this period, vowel–like sounds become more elongated and varied (Hoff, 2001; Ramsdell–Huddock et al., 2019). Glottal sounds may be observed to interrupt these vowel–like productions, and vocal quality becomes higher–pitched and harsh sounding (Nathani et al., 2006). Consonant–like sounds may also be observed, especially consonant approximations formed in the posterior oral cavity (Hoff, 2001; Ramsdell–Huddock et al., 2019). This *cooing and gooing*, as it is commonly called, occurs most often when babies are content or interacting with caregivers (Stark, 1980). Typically, laughter appears near the end of this stage around 16 weeks of age (Hoff, 2001).

The *expansion* or *vocal play* stage of development occurs between 3 to 8 months of age (Hoff, 2001; Nathani et al., 2006; Ramsdell–Huddock et al., 2019). Vocal behavior here is marked by an increase in the variety of consonant–like and vowel–like sounds, which are combined into longer and more complex patterns of production. This *marginal babbling* contains consonant–vowel combinations, albeit with more distortions and slower transitions than an adult model (Goldstein & Schwade, 2008; Nathani et al., 2006; Oller et al., 1999; Ramsdell – Huddock et al., 2019; Stark, 1980). Vocalizations at this stage are also characterized by greater fluctuations in volume, pitch, and movement with yells, whispers, squeals, growls, and frication noises (i.e., partially obstructed airflow; Hoff, 2001; Nathani et al., 2006; Ramsdell–Huddock et al., 2019; Stark, 1980).

From approximately 6 to 10 months of age children begin *canonical babbling* (Hoff, 2001). True syllables emerge. This stage is a major milestone in prelinguistic development as it demonstrates significant advancement toward mature, adult–like speech patterns (Hoff, 2001; Nathani et al., 2006; Oller, 1980; Patten et al, 2014). Transitions between clearly articulated consonant and vowel sounds are more rapid, thus mimicking the timing of natural, adult–like speech (Nathani et al., 2006; Oller, et al., 1999; Ramsdell – Huddock et al., 2019; Stark, 1980). Initially, the syllables produced at this stage are reduplicated with unchanging consonant and vowel sounds as in “baba,” but over time they become more variegated (Hoff, 2001; Nathani et al., 2006; Ramsdell – Huddock et al., 2019). With variegated or nonreduplicated babbling, a wider range of consonant and vowel sounds are combined into a longer series of changing sounds (e.g., “madeepaboo”). Whispering, or voiceless productions, are also observed (Nathani et al., 2006; Ramsdell–Huddock et al., 2019).

Jargon or the *advanced forms* stage occurs between 11 and 18 months of age (Nathani et al., 2006; Ramsdell – Huddock et al., 2019). As the final stage of prelinguistic vocal development, it may precede or coincide with the appearance of first words (Hoff, 2001; Nathani et al., 2006). During this period, vocalizations become more advanced with the appearance of diphthongs and complex syllable shapes including closed syllables, disyllables, and multisyllables (Nathani et al., 2006; Ramsdell – Huddock et al., 2019). Infants also begin imitating the prosody of adult speech (Hoff, 2001; Nathani et al., 2006; Ramsdell – Huddock et al., 2019; Stark, 1980). Consequently, these productions contain true syllables with the sounds, timing, stress, and melodic patterns of speech. Given the close resemblance, caregivers often perceive these utterances as potential words and, with social feedback, can influence the phonological patterns and pragmatic use of these utterances (Goldstein & Schwade, 2008; Lee et al., 2018; Oller et al., 1999).

Early Linguistic Period

Typically, children produce their first words around 12 months of age (range 10 and 15 months), although to start with, these may be rough approximations of the adult model and require multiple trials for acquisition (Goldfield & Reznick, 1990; Hoff, 2001). At this point children are not fully talking, but as a result of the vocal development in the prelinguistic period, they can produce many of the sounds necessary for spoken language (Hoff, 2001).

Initially, the words at this stage of development are often idiosyncratic, but phonetically consistent productions, which are tied to particular contexts or take on specific functions such as gaining attention, making requests, or labeling stimuli in the environment (Hoff, 2001; Oller, 1980). These *protowords,* as they are called, are not based on the words of the native language. For example, "daa–doo" used to request a pacifier. They may, however, be related to the context, such as "ooo–ooo" to label trains. True words, on the other hand, are approximations or imitations of the words in the native language (Bauman–Waengler, 2004; Hoff, 2001). While lexical learning is initially slow, the rate of acquisition rapidly increases after children learn their first word and again after acquiring a repertoire of 50 spoken words (Baumam–Waengler, 2004; Goldfield & Reznick, 1990; Hoff, 2001). In their norming study of vocabulary development, Bates et al. (1994) found a significant positive correlation between chronological age and number of words spoken. Between 8 and 11 months of age, children were producing very few true words; no more than three. By 12 months of age the median number of words spoken was six. A significant increase was observed at 16 months of age where the median number of words was 44. As Hoff (2001) explained, after acquiring first words, children learn an average of 8–11 new words each month, and after achieving a vocabulary of 50 words, the rate of learning increases to 22–37 new words each month (see also Goldfield & Reznick, 1990). Many children then acquire a new word after a single exposure. Bates et al. (1994) also found that by 20 months of age, the median number of words spoken was 170, and at 24 months of age the median number was 311. Just 6 months later vocabulary size increased further with a median of 574 words. Notably, the average rate of acquisition during this period was more than 40 new words per month.

Phonological development continues during the transition from the prelinguistic to linguistic period (Hoff, 2001). At the beginning of the linguistic period, articulation is quite variable (Baumam–Waengler, 2004; Hoff, 2001). First words are typically simple syllable shapes (Hoff, 2001). Single syllables such as consonant–vowel (CV) combinations as in “go,” and reduplicated syllables such as “bye–bye” are common. The phonological composition of this early vocabulary usually includes 2–8 different consonants, including /m, b, d, y, n, w, p, h/. Productions may closely mimic the adult model such as “baby,” or be rougher approximations such as /dɑdi/ for “dog” (Hoff, 2001; Tager–Flushberg et al., 2009). When children acquire a vocabulary of 50 words, around 18–24 months of age, they develop a phonological system, and their speech productions become more consistent (Hoff, 2001). This improved consistency is even apparent within errored productions. They exhibit phonological processes, which change the syllable structure and/or individual sounds in a systematic manner to accommodate their limited articulation abilities at this stage (e.g., /wɑwɑ/ for “water,” /ʌ/ for “up,” /dus/ for “juice,” /mʌni/ for “bunny”; Baumam–Waengler, 2004; Hoff, 2001).

Vocal Deficits in the Prelinguistic and Early Linguistic Periods

By retrospectively examining home videos, researchers have been able to analyze and, in most cases, compare the vocal behavior of individuals later diagnosed with autism spectrum disorder (LDA) and those who are typically developing (TD), during the first two years of life. As previously explained, these prelinguistic vocal behaviors are critical to the later development of speech and language skills (Hoff, 2001; Roche et al., 2018). Understanding the atypical vocal productions associated with autism in early childhood may not only facilitate early detection, diagnosis, and access to treatment, but also aid in the development of specific interventions that address these important precursors to later communication and social behaviors (Maestro et al, 2001).

Birth to 6 Months

Maestro et al. (2002), Brisson et al. (2014), and Chericoni et al. (2016) analyzed vocal productions in LDA and TD infants from birth to 6 months. Although there were no differences in vocal pitch, duration, (Brisson et al., 2014) or rate (Chericoni et al., 2016) between the two groups, Brisson et al. found that the LDA infants produced more simple contours (i.e. melodic changes) and significantly fewer complex contours than the TD infants. Maestro et al. found that the TD infants emitted significantly more vocalizations towards people than did the LDA infants. In a similar study, Zapella et al. (2015) examined the vocal behavior of 18 infants who presented with behaviors characteristic of autism in their second year of life. Only 10 were later diagnosed with autism, while the remaining eight acquired communication abilities consistent with their chronological age. Twelve infants exhibited vocal behavior appropriate for their age. The remaining six, including three of the LDA infants, did not exhibit cooing until 5–6 months of age and their vocalizations were described as having poor melodic changes and more closely resembling those of a TD 2–month–old infant.

6 to 12 Months

The vocal behaviors demonstrated later in the prelinguistic period – sounds, syllables, and prosody – are the same components of typical speech patterns and expressive language (Oller et al., 1999; Patten et al., 2014). As such, canonical babbling is a critical milestone in language acquisition (Hoff, 2001; Patten et al., 2014). According to Chericoni et al., (2016) the vocal trajectories of LDA infants begin to most notably deviate from typical trajectories of development during the 6–12-month age range. Delays in canonical babbling beyond 10 months of age have been reported as possible indicators of later communication deficits (Chericoni et al., 2016; Oller et al., 1999; Roche et al. 2018).

Through retrospective video analysis, LDA infants have been observed to emit significantly fewer vocalizations than TD infants at 6–12 months of age (Chericoni et al., 2016) and 9–12 months of age (Patten et al., 2014). Werner et al. (2000) reported lower rates of vowel sounds per minute in 8–10–month–old LDA infants, and Patten et al., (2014) identified a significantly lower ratio of canonical syllables as compared to TD infants. At 9–12 months of age, TD infants were 17 times more likely to be in the canonical babbling stage than LDA infants (Patten et al., 2014). Conversely, Werner et al. (2000) and Osterling et al. (2002), who examined 8–10 months old infants and 12–month–old infants, respectively, found a higher rate of canonical syllables and/or other vocal behaviors in LDA infants, although these differences were not statistically significant. On the other hand, Maestro et al. (2005) found no significant differences in vocalizations towards objects or people for LDA and TD infants 7–12 months of age.

12 to 24 Months of Age

The *first words* stage is quite salient (Mayo et al., 2013). Caregivers tend to be sensitive to the development of true, spoken words, and any observed delays here are, perhaps, one of the most blatant signs of communication deficits in early childhood (Chericoni et al., 2016; Mayo et al., 2013). When analyzing videos of 12–18–month–old toddlers, Chericoni et al. (2016) found that LDA toddlers produced a lower frequency of first words than TD toddlers, although Maestro et al. (2001) did not observe differences in semantically meaningful vocal behaviors between the two groups until 18–24 months. Similarly, Mars et al. (1998) found that LDA toddlers 12–30 months of age, spoke fewer words and had less verbal imitation than TD toddlers. Werner and Dawson (2005) found that 12–month–old LDA infants, who experienced regression in the first 3 years of life, actually produced more complex babbling and words than TD infants and those with early onset autism (i.e., symptoms present at 1 year of age). By 24 months of age though, the TD infants produced significantly more single words and 2+ word phrases, than either of the LDA groups.

Differences in babbling also persisted at this stage. Patten et al. (2014) reported that TD toddlers 15–18 months of age were six times more likely to be in the canonical babbling stage than LDA toddlers. An overall lower rate of vocalizations at 12–18 months (Chericoni et al., 2016) and 15–18 months (Patten et al., 2014) was observed in LDA toddlers as compared to TD toddlers. Werner and Dawson (2005) found no differences in the frequency of simple babbling among the LDA early onset, LDA with regression, and TD groups at 12 months. However, at 24 months of age, the TD toddlers produced significantly more complex babbling than either of the LDA groups.

Despite observed variations in research outcomes, these studies show that vocal atypicalities for individuals on the autism spectrum are often present in the prelinguistic and/or early linguistic stages of development (Brisson et al., 2014; Chericoni et al., 2016; Maestro et al., 2002; Mars et al., 1998; Oller et al., 1999; Patten et al., 2014; Werner et al., 2000; Zapella et al., 2015). As compared to typically developing children, infants and toddlers who are later diagnosed with autism, exhibit an overall reduced frequency or rate of vocal behavior, are less likely to advance to the canonical babbling stage, and are less likely to produce more complex vocal behavior, such as variegated babbling and true words (Roche et al., 2018). These early, vocal atypicalities may simply be signs of the language impairments inherent to autism, which have yet to fully manifest. However, it is also possible that these vocal delays directly contribute to the social communication deficits apparent later in life (Patten et al., 2014).

Language learning occurs within a social context (Chericoni et al., 2016; Patten et al., 2014). Caregivers attribute meaning to their child's vocalizations, helping them to expand vocabulary, develop syntax, learn new pragmatic functions, and acquire the acoustic features of adult speech (Warlaumont et al., 2014). As such, caregiver feedback is vital to the development of speech, language, and cognition, although caregivers are more inclined to provide such feedback following their child's speech–like vocalizations (e.g., babbling, true words). Warlaumont et al. (2014) describes this phenomenon as a *social feedback loop,* which is based on constructivist theories of cognitive development. In the social feedback loop, the vocal behavior of the child and caregiver are dependent on one another. The child produces a speech–like vocalization, as opposed to crying, laughing, or vegetative sound, and the caregiver responds immediately and positively, which then increases the probability of similar, subsequent vocalizations from the child. If infants and toddlers, who are later diagnosed with autism, emit fewer speech sounds, they may receive less feedback and social engagement from caregivers, which further limits the language they are exposed to, making it more challenging to learn functional social communication and become proficient vocal communicators (Chericoni et al., 2016; Patten et al., 2014; Warlaumont et al., 2014).

Nature of Vocal Communication Deficits in Autism

While the social communication deficits associated with autism may be further exacerbated by the weakened contingency between child–caregiver vocalizations (Patten et al., 2014; Warlaumont et al., 2014), it is necessary to identify the nature of these initial speech delays/deficits and design effective interventions that address the underlying cause (Shriberg, 2010). Beiting (2022) recognized that the complexities of autism, which may include deficits in areas of attention, cognition, communication, imitation, social motivation, and sensory–motor abilities, make possible any number of variables to impede the acquisition of spoken language. Yet for minimally vocal children in particular, motor speech disorders, namely childhood apraxia of speech, is often considered causally associated with the observed vocal communication deficits (Adams, 1998; Beiting, 2022; Chenausky et al., 2019; Gernsbacher et al., 2008; Shriberg, 2010; Tierney et al., 2015).

Childhood Apraxia of Speech

In a survey of 132 speech–language pathologists, Dawson (2010) reported that speech–language pathologists working in early intervention suspect, on average, that 1 out of every 6 children on the autism spectrum also has apraxia of speech. Of the 16% of children suspected, half were described as nonverbal. As Dawson (2010) explained, children on the autism spectrum may exhibit reduced motivation for social communication, as well as deficits in many areas of language, but they may also lack sufficient motor coordination to produce speech sounds and sequence them together for fluent, intelligible vocal communication. Unfortunately, diagnosing childhood apraxia of speech is quite difficult, especially in children with limited vocal output.

Adams (1998) compared the motor speech behaviors and nonspeech oral movements of four children on the autism spectrum and four typically developing children, 6–11 years of age. The participants were matched for gender, age, and ethnicity. The *Kaufman Speech Praxis Test* was administered to assess 11 oral movements, 70 simple phonemic/syllabic productions, and 90 complex phonemic/syllabic productions. All tasks required imitation of the examiner. The participants in the control group earned perfect scores on the oral movement tasks and simple phonemic imitation. Three out of four earned a perfect score on the complex phonemic imitation. Statistically significant differences between the groups were reported for the oral movements and the complex phonemic/syllabic productions. In particular, the children on the autism spectrum had difficulty imitating tongue tip elevation and alternating lip rounding and retraction. Errors on the simple phonemic imitation task included prevocalic voicing of voiceless phonemes and substitutions in vowel–consonant–vowel (VCV) and vowel–consonant–vowel–consonant (VCVC) syllables. The greatest differences between the groups were observed on complex phonemic imitation. The children on the autism spectrum exhibited oral scanning/groping, syllable deletion, consonant and vowel distortions, cluster reduction, fronting, and gliding. To summarize, the children on the autism spectrum had difficulty with nonspeech oral movements and exhibited articulation errors consistent with motor speech impairments, such as apraxia and dysarthria.

Marili et al. (2004, as cited in Velleman et al., 2009) surveyed parents of 40 children on the autism spectrum regarding their speech behaviors. The children ranged in age from 22 months to 22 years. Parent reported speech behaviors were consistent with motor speech impairments for 60% of the participants. More specifically, 12.5% of the participants reportedly exhibited symptoms of apraxia, 10% exhibited characteristics of dysarthria, and 37.5% presented with speech patterns that could be associated with either motor–related condition.

Using standardized assessments and video coding, Chenausky et al. (2019) evaluated 54 children with low verbal abilities (LV; i.e., spontaneous phrase speech, but not verbally fluent) and minimal verbal abilities (MV; i.e., no spontaneous phrase speech) to examine the degree of motor speech impairments in children on the autism spectrum with severe vocal–verbal communication deficits. The participants ranged in age from 4 years, 4 months to 18 years, 10 months. Expressive language was measured as the number of different words spoken during testing. As explained, spontaneous vocal words capture aspects of both speech and language, demonstrate communicative intent, and are not necessarily limited by imitation abilities. The participants were divided into four groups, based on 1) their performance on the *Kaufman Speech Praxis Test* and 2) the presence of speech errors consistent with childhood apraxia of speech, such as vowel errors, consonant distortions, stress errors, and groping.

Group 1 included 12 children, six LV children and six MV children, whose speech was within normal limits for their chronological age (Chenausky et al., 2019). While these children exhibited mild and infrequent abnormalities, such as consonant distortions, voicing, and prosodic errors, their articulation errors were not clinically significant. In total, the participants in group 1 produced a mean of 1.8 speech abnormalities (range 0–9). The number of different words spoken during assessment ranged from 9–229 with an average of 101.8.

Group 2 consisted of 16 children, two LV and 14 MV, who demonstrated vowel reduction, hypernasality, placement errors, and /r/ distortions (Chenausky et al., 2019). Speech productions in this group were described as “underarticulated” or “mumbled” (p. 5), but not characteristic of childhood apraxia of speech. According to Chenausky et al. (2019), this group was likely comprised of mild, but miscellaneous speech impairments such as phonological disorders, developmental errors, or dysarthria. The mean number of speech abnormalities was 5.2 (range 1–14), and they produced 4–211 different words with an average of 62.3.

Group 3 included 13 MV children who presented as severely disordered (Chenausky et al., 2019). The most common errors included consonant distortions, syllable segregation, addition of phonemes, and errors in stress, voicing, nasality, and vowels. An average of 7.7 speech abnormalities were reported (range 1–12). At least five signs of childhood apraxia of speech were present, and the children exhibited significant difficulty with motor planning and programming for speech productions. Thus, the participants in group 3 were suspected to have childhood apraxia of speech. The number of different words spoken ranged from 0–45 with an average of 14.2.

Finally, group 4 included 13 MV children who could not be formally evaluated due to their limited vocal imitation (Chenausky et al, 2019). Three of the participants did not vocalize during the assessment, five emitted nonspeech vocalizations, and, while the remaining five vocalized on request, their vocal behaviors were described as having a breathy, harsh vocal quality and/or limited to imitation of simple syllables. These children produced 0–34 different words during the assessment with an average of 7.5.

The speech and language skills of the children in the study were quite diverse (Chenausky et al., 2019). Even from a sample of children on the autism spectrum with minimal and low verbal abilities, some did not emit any vocalizations, others produced over 200 words, some had age–appropriate speech, and others had severely impaired articulation. When the children were categorized according to their speech abilities, statistically significant differences in the mean number of speech abnormalities were seen across the groups. Interestingly, performance on the speech portion of the *Kaufman Speech Praxis Test* significantly predicted the number of different words produced by the participants in groups 3 and 4. The children in group 3 were the only participants who presented with characteristics of childhood apraxia of speech, but further assessment is needed to confirm such a diagnosis. If these children do indeed have a motor speech impairment, then the question remains whether apraxia is a co-occurring condition or an inextricable subphenotype of autism.

Gernsbacher et al. (2008) examined the relationship between the motor skills of infants and toddlers and their later speech fluency in middle childhood and adolescence. The study included 115 children on the autism spectrum and 44 typically developing children. Parent interviews and home video reviews were conducted to gather historical information on children's early oral and manual motor abilities. Statistically significant differences between the groups were reported. Parents reported their children's current vocal speech abilities, based on school or other professional assessment, and the children on the autism spectrum were classified as highly fluent (i.e., speech within one year of chronological age and rarely unintelligible to unfamiliar people), moderately fluent (i.e., speech 1–3 years below chronological age characterized by noticeable mispronunciations, atypical rate, pitch, or volume, occasional echolalia, and sometimes unintelligible to unfamiliar people) and minimally fluent (i.e., unable to produce 4+ word utterances after six years of age or unable to produce recognizable words at any age). Minimally fluent children exhibited greater deficits in oral motor performance in early childhood than the highly fluent children. Because oral motor movements underlie vocal speech production, the authors contended that the two coincide in minimally fluent children restricting phonetic repertoires to vowels, voiced consonants, and bilabial plosives.

Tierney et al. (2015) evaluated the diagnostic accuracy of the *Checklist for Autism Spectrum Disorder*. Specifically, they were interested in its capability of distinguishing between autism spectrum disorder and childhood apraxia of speech. The study included 30 children who presented with concerns regarding speech, language, or autism. A developmental pediatrician evaluated the children for autism using the *Checklist for Autism Spectrum Disorder*, *Childhood Autism Rating Scale–2*, a records review, and a checklist based on the diagnostic criteria detailed in the *Diagnostic and Statistical Manual of Mental Disorders*. Those children meeting the diagnostic criteria and earning a score of 30+ on the *Childhood Autism Rating Scale–2* were diagnosed with autism spectrum disorder. A speech–language pathologist evaluated the children for apraxia of speech using the *Kaufman Speech Praxis Test*, *Preschool Language Scale–5* or the *Clinical Evaluation of Language Fundamentals*, and parental interview. Those children who exhibited speech patterns characteristic of apraxia of speech based on observation and performance on the *Kaufman Speech Praxis Test,* were diagnosed accordingly.

Tierney et al. (2015) found that 63.6% of the children diagnosed with autism also received a diagnosis of childhood apraxia of speech, and 36.8% of children diagnosed with apraxia also had autism. Out of the 30 participants, 40% were diagnosed with apraxia but did not meet criteria for autism, 13.3% were diagnosed with autism but not apraxia, 23.3% had both conditions, and 23.3% had neither. Diagnostic accuracy of the *Checklist for Autism Spectrum Disorder* was 100% for children who did not have apraxia and 94.7% for those with apraxia. Given the high concomitance of autism spectrum disorder and childhood apraxia of speech observed in this sample, the authors recommended that when children are diagnosed with one of these conditions, they also be evaluated for the other. Proper identification and diagnosis is necessary for the most efficacious treatment.

Velleman et al. (2009) evaluated 10 children on the autism spectrum, ages 4 – 6 years old, for motor speech disorders. All children had expressive vocabularies of at least 50 words, including alternative/augmentative forms of communication, although at least10 of those words were communicated vocally. Most children were combining words into phrases. Speech samples were collected and analyzed for maximum phonation, pitch, duration, volume, length of pauses, length of vowels, voicing, and rate. Prosodic abnormalities, as compared to typically developing children, were reported in 80% of the participants including lower pitch, higher resonance frequency, lower phonation duration, more variable stress, and differences in intonation. The Global Motor Control, Focal Oromotor, and Sequencing subtests of the *Verbal Motor Production Assessment for Children* were also administered. All children scored below the mean for their chronological age on some portion of the *Verbal Motor Production Assessment for Children*. Global Motor Control scores ranged from 90–100 with a mean of 97. Five of the children were within normal limits, while the remaining five earned severity ratings of mild to moderate. Scores on the Focal Oromotor Control subtest ranged from 48 – 91 with a mean of 75.7. Three children were within normal limits, one earned a severity rating of moderate, and the remaining six presented with severe deficits. Sequencing subtest scores ranged from 39 – 98 with a mean of 72. Five children fell within normal limits, two were rated as mild, one moderate, and two were severely impaired. In some ways, the speech patterns of the children on the autism spectrum were like those of typically developing children, and in other ways their speech was similar to children with childhood apraxia of speech (who do not have autism). Consequently, the authors could not confirm the presence of childhood apraxia of speech for any of the participants.

While these studies suggest that autism spectrum disorder and childhood apraxia of speech may indeed co–occur at a relatively high rate, the prevalence of these concomitant conditions may be lower than expected (Shriberg, 2011). According to Shriberg et al. (2011) childhood apraxia of speech is often suspected as a cause of limited or absent vocal communication in children on the autism spectrum, and, in children with some vocal communication, is believed to contribute to abnormalities in speech and prosody. Three main reasons motivate these suspicions: motor impairments, common genetic origins, and phenotypic similarity. First, because many children on the autism spectrum present with praxis deficits that affect imitation and overall gross and fine motor function, (e.g., Dziuk et al., 2007; Gernsbacher et al., 2008; Mostofsky et al., 2006) many professionals generalize these deficits to the speech domain. Yet, Shriberg et al. (2011) argue that the neural substrates underlying praxis conditions, such as limb apraxia, ideational apraxia, and buccofacial apraxia, are not the same as those affecting the speech system. Secondly, both autism and childhood apraxia of speech have been associated with a common gene, *FOXP2*. However, the distinctly different prevalence rates of autism and childhood apraxia of speech, with autism occurring at a considerably higher rate, cast doubt on a common genetic origin to explain comorbidity. Finally, concomitant apraxia is suspected because the speech patterns of children with autism are like those of children with apraxia of speech. But Shriberg et al. (2011) contend that the inconsistent descriptions and heterogeneous findings across the research, prevent direct comparisons and generality (see also Beiting, 2022). Instead, they purport that the observed speech deficits in the autism population are better explained by other speech sound disorders.

Speech Sound Disorders

Speech sound disorders is an umbrella term for a variety of impairments in vocal speech production, defined by the American Speech–Language Hearing Association (ASHA) as, “any difficulty or combination of difficulties with perception, motor production, or phonological representation of speech sounds and speech segments—including phonotactic rules governing permissible speech sound sequences in a language" (n.d.–c. "Overview"). ASHA divides speech sound disorders into organic and functional types. Organic speech sound disorders are developmental or acquired deficits resulting from motor/neurological impairments (e.g., apraxia, dysarthria), structural abnormalities (e.g., cleft palate, orofacial anomalies, anatomical or physiological defects due to surgery or trauma), or sensory/perceptual disorders (e.g., hearing loss). Functional speech sound disorders are idiopathic, though they may be related to motor production (e.g., articulation) or the linguistic aspects of speech (e.g., phonology).

Currently, there is little agreement on standardized measures, subtypes, and severity ratings of speech sound disorders (Shriberg et al., 2011, 2019). Shriberg et al. (2011, 2019) proposed a classification system that divides speech sound disorders into idiopathic (i.e., functional), or those of unknown origin, and complex neurodevelopmental disorders of genetic or environmental origins (i.e., organic). These are further divided into six subtypes.

Speech delay presents as a pattern of developmentally inappropriate substitutions, deletions, and/or distortions in children 3–9 years of age (i.e., phonological disorder; Shriberg et al., 2011, 2019). Prosody and voice remain intact. Speech intelligibility may be mildly to severely impaired, and other domains such as language and reading may be negatively affected. As the name implies, delays will often resolve with treatment (Shriberg, 2010). Population prevalence estimates of speech delays range from approximately 15% at 3 years of age to 4% at 6–8 years of age (Shriberg et al., 2011). Comparable rates have been reported for children on the autism spectrum, with estimates at 12%–15% for children 3–9 years of age (Cleland et al., 2010; Shriberg et al., 2011).

Speech errors most often include distortions of sibilants (e.g., /s, z, ʃ /) and rhotics (e.g., /r/ and vocalic /r/) in children 6–9 years of age (Shriberg et al., 2011, 2019). Atypicalities in prosody and/or voice are not observed. Generally, speech errors do not impact other areas of development/functioning and have little effect on overall speech intelligibility (Shriberg, 2010; Shriberg et al., 2011). Data on prevalence are more limited, but estimated to be approximately 8% among 8–year–old children (Shriberg et al., 2011). Shriberg et al. (2011) reported a much higher prevalence in the autism population, with 31.8% of high functioning, vocal–verbal children, 6–7 years of age, presenting with speech errors. Similarly, Rapin et al. (2009) reported that 24% of children on the autism spectrum, ages 3–9 years, exhibited characteristics of speech delay or speech errors.

Persistent Speech Disorders are speech delays or speech errors that persist past 9 years of age and, in some cases, into adulthood (Shriberg et al., 2011). Population prevalence data indicate that 2.4% – 3.9% of individuals 9 years of age to adult have speech errors. Cleland et al. (2010) reported that 33% of children on the autism spectrum, 5–13 years of age, presented with speech errors or persistent speech disorders (see also Shriberg et al., 2011). Shriberg et al. (2001) also found that 33% of individuals 10–49 years of age, described as having high functioning autism, presented with persistent speech disorders. In 2019, Shriberg et al. found that 16.7% of the 42 participants on the autism spectrum had speech delay or a persistent speech delay while none of the participants presented with speech errors or persistent speech errors.

Motor Speech Disorders comprise the final three subtypes – *Apraxia, Dysarthria,* and *Not Otherwise Specified.* All are characterized by deletions, substitutions, and distortions in speech, as well as significant impairments in prosody and voice, severely reducing intelligibility and affecting individuals of all ages. Apraxia of speech is an impairment in motor planning and/or programming spatiotemporal parameters of movement which affects the precision and consistency of the articulatory sequences for speech production (ASHA, n.d.–a.; Shriberg et al., 2011). Again, data on population prevalence are limited, but rates of apraxia are estimated to be less than 0.1% (Shriberg et al., 2011). Of the 46 children on the autism spectrum analyzed by Shriberg et al. (2019), none presented with apraxia of speech. Dysarthria is the result of neuromuscular weakness in one or more muscles necessary for respiration or articulation that interferes with speech production and typically results in slurring, mumbling, choppiness, increased or decreased rate, and limited movement (ASHA, n.d.–b.; Shriberg et al., 2011). None of the children on the autism spectrum evaluated by Shriberg et al. (2019) presented with the characteristics of dysarthria. Finally, Not Otherwise Specified is a proposed subtype of motor speech disorders in need of additional research (Shriberg et al., 2019). In theory, this classification would encompass speakers who do not meet diagnostic criteria for apraxia or dysarthria but show symptoms of a motor speech impairment, such as delays in precise, stable articulation (Shriberg et al., 2011, 2019). Population prevalence is estimated to be 4.3%–16.46%, although data are limited (Vick et al., 2014). According to Vick et al. (2014), those with Motor Speech Disorders–Not Otherwise Specified are at the greatest risk for persistent speech disorders.

Indeed, many children on the autism spectrum exhibit speech patterns commensurate with their same–aged peers (Kjelgaard & Tager–Flusberg, 2001; Shriberg et al., 2019), yet those with speech sound disorders have been documented to exhibit phonological processes and articulation errors past the typical age of suppression, including substitutions, omissions, gliding, cluster reduction, and final consonant deletion (Aravamudhan & Awasthi, 2021; Cleland et al., 2010). To summarize, speech delays and speech errors may be more common than motor speech disorders (Shriberg et al., 2011, 2019). These findings are in contrast with commonly held beliefs about the nature of speech sound disorders in autism. However, it is important to note, that these findings are based on the speech profiles of children on the autism spectrum with relatively high verbal abilities (Beiting, 2022; Shriberg et al., 2011). It is, in fact, possible that limited/absent speech–like vocalizations, as is the case with minimally verbal children on the autism spectrum, may be the result of a severe underlying motor impairment. But, apraxia of speech, and other motor speech disorders, are diagnosed based on the characteristics of vocal behaviors (Chenausky et al., 2019). Therefore, it is impossible to evaluate motor function without sufficient vocal output. Nevertheless, speech sound disorders further inhibit the acquisition of spoken language for children who already present with social communication deficits and learning challenges, and thus necessitates effective assessment and intervention (Beiting, 2022; Cleland et al., 2010).

Treatment of Speech Sound Disorders in Minimally Vocal–Verbal Children on the Autism Spectrum

Teaching vocal communication to minimally verbal individuals on the autism spectrum seems to receive little attention. Koegel et al. (2020) conducted a systematic review of scholarly articles that targeted vocal–verbal expressive communication in minimally verbal or nonverbal individuals on the autism spectrum. Only 31 articles, published between 1960–2018, met the criteria. Treatment of speech sound disorders, in particular, also seems to be a relatively understudied area. A systematic literature review by Beiting (2022) found only 14 studies with publication dates from 1979–2021, that focused on the assessment or treatment of speech sound disorders in children on the autism spectrum. In their review of verbal behavior research in autism, DeSouza et al. (2017) identified 172 publications between 2001–2017, that had at least one verbal operant as the dependent variable. Most studies targeted mands (n=91), followed by tacts (n=56), intraverbals (n=40), and lastly echoics, which were targeted in only four studies. As Chenausky et al. (2016) explained, teaching minimally verbal children independent and functional language is certainly important, but it is equally important that we teach vocal speech. Improving speech intelligibility fosters social communication, as it helps establish verbal responses that will be acknowledged and responded to by others, so that individuals with autism may become full, active communication partners (Brady et al., 2016; Chenausky et al., 2016).

Auditory–Motor Mapping Training

Auditory Motor Mapping Training (AMMT) is an intonation–based approach to teach speech production (Wan et al., 2011). It associates the auditory stimuli with motor speech movements and activates neural pathways of the brain, strengthening connections between auditory and motor functions. While singing a bisyllabic target word or phrase, the therapist taps electronic drums that are tuned to the same pitch as the vocal targets and the child learns to imitate. The treatment is derived from intonation–based techniques used to improve speech production in left–hemisphere stroke patients (Chenausky et al., 2017). During the AMMT sessions, the therapist introduces an associated picture and the vocal target by singing it at a rate of one syllable per second and simultaneously tapping the drums in an alternating pattern (Chenausky et al., 2016, 2017; Wan et al., 2011). Next, the therapist and child produce the vocal target in unison. On subsequent trials, the therapist produces the target with the child but stops after the first syllable so that the child completes the last syllable independently. Finally, the therapist models the vocal target and asks the child to imitate.

Wan et al. (2011) evaluated AMMT in teaching vocal behaviors to children with no functional vocal–verbal abilities. The study included six children on the autism spectrum, ages

5–9 years, who had no intelligible words. Speech therapy had primarily focused on teaching augmentative/alternative modes of communication including the *Picture Exchange Communication System*, gestures, signs, and a speech generating device. Although imitation was part of the training, the probe assessments only presented the associated pictures to evoke the vocalization. Probe assessments were conducted every 5–10 sessions before, during, and after treatment and evaluated trained and untrained vocalizations. Effects of AMMT were evaluated in a single–subject AB design. All participants showed improvements above baseline for both trained and untrained targets in number of vocal approximations and exact productions. Vocal approximations were deemed correct if they contained a correct vowel sound and a consonant with at least two out of three correct features (i.e., place, manner, or voicing). Exact productions improved from 0% correct at baseline to an average of 29% post–training (range 8%–71%).

Chenausky et al. (2016) compared the effects of AMMT and Speech Repetition Therapy (SRT). SRT was described as a conventional speech treatment which emphasizes imitation and repetition of vocal targets. Vocalizations are typically produced at a normal speaking rate and the therapy does not include intonation changes or tapping of drums as in AMMT. In total, 30 minimally verbal children on the autism spectrum (i.e., less than 20 intelligible words and no productive syntax), 3 years, 5 months to 9 years, 8 months of age, participated in the study. Fourteen children, matched for chronological age, mental age, and speech sound imitation, were equally divided into two groups: AMMT and SRT. The remaining nine children received AMMT. Vocal targets were individualized for each participant, but included 15 bisyllabic, high frequency words or phrases. Probe assessments evaluated the child's imitation of trained and untrained vocal targets across three dependent measures: number of syllables approximated, number of consonants correctly produced, and number of vowels correctly produced per word.

Vocal responses were scored in the same manner as Wan et al., (2011) except that the vowel sound was considered correct if it was in the same class as the target (e.g., tongue height, stress).

According to the results, participants in the AMMT group showed significant improvements in the percentage of syllables approximated, the percentage of consonants correct, and the percentage of vowels correct (Chenausky et al., 2016). Although there was no difference in vowel production between the matched groups, the AMMT group showed greater gains in the percentage of syllables approximated and the percentage of consonants correctly produced. All 23 participants who received AMMT responded to the treatment, and, in the matched groups, a significantly higher number of AMMT participants responded to treatment as compared to the SRT participants. Improvements were also reported across trained and untrained targets.

Chenausky et al. (2017) also compared the effects of AMMT and SRT on vocal behavior. The study included four children on the autism spectrum between the ages of 4 years, 1 month and 6 years, 7 months. Two children were described as minimally verbal (i.e., less than 20 spoken words and no productive syntax) while the other two were reportedly "more verbal". Specifically, these verbal children had a spoken vocabulary of 10–50 words and were combining words into phrases with a mean length of utterance ranging from 1.1 – 2.4. All children were able to imitate at least two speech sounds, participate in activities at the table for 15+ min, follow 1–step directions, and imitate simple oral and gross motor movements. To compare AMMT and SRT across verbal abilities, one minimally verbal child and one verbal child received AMMT and the other two received SRT. As in Chenausky et al. (2016) vocal targets included 15 bisyllabic words/phrases. Targets for the verbal children included 16 bisyllablic or trisyllabic words/phrases. Methods were like those of Chenausky et al. (2016). However, the final step of the training phase included an echoic to intraverbal transfer procedure. For example, after learning to imitate "bubbles," the therapist would say, "It's fun to blow ___" and pause for the child to independently say "bubbles".

In both the verbal and minimally verbal children, AMMT produced statistically significant gains in number of correct consonants, correct vowels, and correct syllables per target as compared to SRT (Chenausky et al., 2017). When improvements were averaged across participants, over both trained and untrained targets, increases were reported. An exception to this was among the verbal participants in the number of correct syllables per target. The AMMT participant improved by an adjusted mean of 0.5 syllables correct per trained target and 0.9 syllables correct per untrained target. On the other hand, the SRT participant improved by an adjusted mean of 0.1 syllables correct per trained target and decreased by 0.3 syllables correct per untrained target.

Chenausky et al. (2016) proposed three reasons that AMMT may be more effective than SRT. First, the musical activities of AMMT may be preferred by participants, and thus increase motivation for participation and engagement. Secondly, AMMT may activate the auditory and motor regions of the brain and connect neural pathways, thus facilitating the acquisition of spoken language. Finally, based on the suspicion that many children on the autism spectrum have co–occurring apraxia of speech, AMMT incorporates many elements of successful interventions for motor speech impairment including imitation, reduced rate, and simultaneous production.

While Wan et al. (2011) and Chenausky et al. (2016, 2017) showed improvements in speech production following AMMT, the results should be interpreted with caution. In each of these studies, improvements were seen in both trained and untrained targets. The authors attributed these findings to the generality of the treatment, but the lack of functional control in the experimental design, suggests other variables (e.g., participant characteristics, complexity and/or familiarity with the vocal targets) may be responsible for the observed changes.

Technology–Based Interventions

Computer Programs. Using a multiple baseline design across participants, Beiting and Maas (2021) evaluated the effects of a treatment package, Autism–Centered Therapy for Childhood Apraxia of Speech (ACT4CAS), on the speech productions of three children on the autism spectrum with co–occurring childhood apraxia of speech. The participants ranged in age from 4 years, 6 months to 7 years, 3 months, and all were described as minimally verbal, in that they emitted less than 30 different words during a spontaneous language sample. Standard scores on the *Goldman Fristoe Test of Articulation – 2*, were less than 40 for Participants B and C, indicating severe deficits in speech production. Participant A was unable to complete the standardized assessment. Having the most limited speech of all the participants, his phonetic inventory included five simple vowel sounds and three early consonants.

Two targets were selected for Participant A and four targets each were selected for Participants B and C (Beiting & Maas, 2021). Treatment began with 10 min of drill–based practice. Targets were presented in practice sets on an iPad with prerecorded videos and static images. The practice set included different speakers modeling the vocal target (e.g., "I'm sleepy"), and the pictures and videos were used to express the concept associated with the target word/phrase. For example, the audio "I'm sleepy" was accompanied by a picture of a child sleeping and a video of a dog sleeping. A predetermined prompt hierarchy including physical prompts, model prompts, and verbal prompts, was implemented to promote accurate imitation of the target word/phrase. The next phase of treatment was 10 min of play–based practice. Approximately four activities were planned around the target word/phase. For example, the target word “eat” included a snack, play with pretend food, and puppet play. Caregivers were encouraged to participate, and the adults modeled the target word/phrase, but did not directly request imitation. Following the play–based practice, the participant was provided a short break, and the second vocal target was addressed with the same procedures. Imitation probes were conducted at the start of each session. Accuracy of production was measured according to the correct number of syllables, correct sequence of sounds, and correct prosody including stress and transitions between sounds, across probes conducted at the start of each session. None of the participants improved production of any vocal targets in the treatment condition, but Participant B showed slight improvements of two targets at the follow–up measures conducted at one week and three weeks post–treatment. As such, this treatment does not appear to be effective for minimally verbal children on the autism spectrum with co-occurring childhood apraxia of speech.

Chen et al. (2019) compared the effectiveness of a 3–D virtual tutor to a real human face for teaching Mandarin speech to children on the autism spectrum. The 3–D virtual tutor was a computer–generated, audio–visual animation of a human head with the face, lips, tongue, jaw, and nasopharyngeal wall visible in both frontal and profile views. Realistic visual speech was modeled through the animated articulatory movements and airflow changes visible on the virtual tutor. The speech of the 3–D virtual tutor was not synthesized but was recorded from the real human face tutor (a female language teacher) to provide a more natural auditory model. Recordings of the real human face tutor were similar with movements, frontal and profile views, and naturalistic audio all presented on the computer screen.

Participants were 22 children on the autism spectrum who were living in Mandarin–speaking communities (Chen et al., 2019). The children were 3–6 years of age and described as “low functioning”. In pretest and posttest conditions, participants were asked to imitate 16 different Mandarin syllables. Using a 5–point Likert scale, the vowels and consonants in each syllable were scored separately ranging from 1, *completely incorrect*, to 5 *completely correct*. Participants were equally divided into two groups: 3–D virtual tutor and real human face. The groups did not differ significantly in mean chronological age or pretest performance. During training, participants were instructed to watch the videos and imitate the tutor. The 16 syllables were presented four times per session across three sessions. Both groups showed a statistically significant improvement in consonant and vowel imitation after training. Overall, in the 3–D virtual tutor group, consonant production increased from a mean of 3.39 in pretest to 3.99 in posttest. Vowel production increased from 3.61 in pretest to 4.19 in posttest. In the human face tutor group, consonant production increased from an average score of 3.42 in pretest to 3.79 in posttest. Vowel production increased form a mean of 3.63 to 4.01. The 3–D virtual tutor group showed statistically greater gains in articulation of all targeted consonant sounds and two vowel sounds. Consequently, computer generated, virtual instruction may be a viable treatment option for improving imitative speech in children on the autism spectrum, especially in areas where therapy services are limited.

Using a simultaneous treatment design, Bernard–Opitz et al. (1999) compared the use of a computer program, SpeechViewer, with traditional play interaction for increasing vocal imitation in children on the autism spectrum. The SpeechViewer provided visual feedback through various graphic displays in response to sound productions. It included four subprograms each focusing on a different speech skill: sound awareness, loudness awareness, voicing onset awareness, and vowel accuracy skill building. Each program targeted specific sounds/words through fun, voice–activated, children’s games. For example, the voicing onset awareness subprogram incorporated a train to teach the CV syllable /tʌ/. When the child voiced, the train would move down the track with the goal of reaching the station house.

The participants included 10 children on the autism spectrum, ages 3–7 years. All children were described as “nonverbal” (Bernard–Opitz et al.,1999). When presented with speech sounds, five children could not imitate. The remaining five children attempted to imitate, although inconsistently. All were able to imitate simple gross motor actions (e.g., clap hands, tap the table). Prior to each treatment condition the parent or therapist modeled the target sound, attempting to evoke imitation. Approximations or correct imitations were scored as correct. New sounds were introduced once a child obtained 80% correct.

In the SpeechViewer condition, the children completed all subprograms with a parent or therapist (Bernard–Opitz et al.,1999). The adult modeled the sound and pointed to the screen. When the child imitated, the computer would display color changes, movement, or other effects. For example, after imitating /bʌ/, a balloon would expand on the screen. During the traditional play interaction, the adult would engage the child with a toy and model the target sound. Imitations or approximations were followed by verbal praise and play interaction. For example, after imitating /bʌ/ the adult would praise the child and bounce the ball playfully. Vocal imitations improved across both conditions, although gains and rate of learning were significantly greater in the SpeechViewer condition. The authors suggest that the effects of the SpeechViewer may be due to the visual feedback or the novelty of the computer games.

Auditory Match–to–Sample. Auditory discrimination is an important precursor to vocal imitation and more advanced linguistic behaviors (Brown, 2005; Du et al., 2017). Learning to listen and match auditory stimuli may facilitate discrimination of speech sounds and imitative responses (Choi et al., 2015; Du et al., 2017). As such, several authors have evaluated the effect of auditory match–to–sample tasks on echoic repertoires in children on the autism spectrum using a multiple probe design.

Brown (2005) taught five preschool children, ages 4–5 years, to match auditory stimuli. Participants A, B, D, and E had a diagnosis of autism spectrum disorder. According to the results of Participant C's assessment, he was described as having a possible multisystem developmental delay or autism. Results of the *Preschool Inventory of Repertoires for Kindergarten* indicated that the participants did not emit any echoic responses following a vocal model. For the auditory match–to–sample procedures, the auditory stimuli were presented on BIGmack® buttons, which are single message speech–generating devices that record and playback up to 20 s of audio. Three buttons were arranged in a triangular shape on the table with one button at the top, and two buttons positioned side–by–side just underneath. The instructor and participant sat on opposite sides of the table. The instructor was seated near the top button which presented the target stimulus. The participant was seated near the two bottom buttons. One of the bottom buttons contained a match of the target stimulus while the other contained the non–exemplar. To present the auditory matching task, the instructor pressed the top button, and each of the bottom buttons with instructions such as "Your turn," or "Which one matches mine?". The participant was to press the bottom button with the same auditory stimulus as the top button. Preferred stimuli were delivered contingent on correct responses while an error correction procedure was implemented following incorrect responses. Generalized reinforcers varied by student but included vocal praise, tickles, or access to a preferred item.

The auditory matching procedures were divided into six phases with progressively finer discriminations (Brown, 2005). Matching began with environmental sounds (e.g., baby crying, cat meowing) in the presence of no sound, followed by white noise and other environmental sounds. When students learned to match sounds, subsequent phases required matching words in the presence of environmental sounds, nonwords (i.e., babbling), and eventually other, true words. Pre and post experimental probes were conducted for two echoic–mands and two echoic tacts. Twenty probes were conducted for each of the four items resulting in a total of 80 probes. The target items selected for the echoic–mand probes were preferred and a deprivation condition was created prior such that the items were only available during the probes. Although the participant could see the preferred item, it could only be accessed with a full echoic response. After obtaining the participant's attention, the instructor modeled the vocal mand and if the participant imitated the model, the preferred stimulus was delivered. Incorrect responses were briefly ignored before the next trial was presented. The same procedures were used for the echoic–tact probes, but a generalized reinforcer, such as praise, was delivered contingent on full echoic responses.

The dependent measures included full echoics (i.e., complete point–to–point correspondence with the model), partial echoics (i.e., partial point–to–point correspondence with the vocal model), and other vocalizations (i.e., a vocal response that followed the model but without correspondence; Brown, 2005). During the pre–experimental probes, Participants A, B, C, and D did not imitate the target vocalizations. More specifically, they did not emit any vocal responses (i.e., full or partial echoic or other vocal response) following the vocal antecedent. Participant E did produce three vocalizations during the pre–experimental probes, although none corresponded with the model (i.e., no full or partial echoics). After mastering the final phase of treatment, during the post–experimental probes, Participant A still did not emit any full echoic responses but produced 20 partial echoics and 32 other vocalizations. Participant B acquired 1 full echoic and 79 partial echoics. Participant C acquired 5 full echoics and 75 partial echoics.

Participant D did not emit any full or partial echoics but produced 3 other vocalizations. Finally, Participant E acquired 58 full echoic responses and 22 partial echoic responses. All participants required 260 – 600 trials to achieve generalized auditory matching (i.e., 100% correct responding on the first session with novel auditory stimuli) and four out of five participants showed significant improvements in echoic responses as a result of the auditory matching treatment.

A second experiment was conducted to evaluate the effects of a generalized auditory word match–to–sample on inexact echoic repertoires (Brown, 2005). The study included four preschool children ages 4–5 years. Two children were diagnosed with autism spectrum disorder and the remaining two did not carry a specific diagnosis. While language and other adaptive skills varied across the participants, they all communicated vocally using phrases and sentences. Varying degrees of articulation impairments were described, but all participants demonstrated inexact echoic repertoires characterized by poor point–to–point correspondence with the vocal antecedent (e.g., "bouse" following presentation of "house," "fawa" following presentation of "flower," "aboons" following presentation of "balloons"). The materials and overall procedures for the auditory matching tasks were the same as those in the first experiment, but the phases were different. Matching began with environmental sounds in the presence of other environmental sounds. Subsequent phases included matching non–words in the presence of non–words, words in the presence of words (with identical voices), and words in the presence of words with varied voices.

The four vocal targets selected for the probe conditions were based on errors observed during the pre–assessments (Brown, 2005). As in the first experiment, 20 echoic probes were conducted for each of the four target vocalizations for a total of 80 probes. During the pre–experimental probes, all participants emitted inexact echoic responses following the vocal model.

After acquiring a generalized auditory word match–to–sample repertoire, all participants showed significant improvements in echoic responses. In the post–experimental probes, Participant F emitted 75 full echoic responses. Her standard scores on the *Goldman–Fristoe 2 Test of Articulation* increased from 70 at the pre–assessment to 107 at post–assessment. Participant G showed 78 full echoic responses during the post–experimental probes and scores on the *Goldman–Fristoe 2 Test of Articulation* increased from 78 at pre–assessment to 112 at post–assessment. Participant H produced 68 full echoic responses and improved scores on the *Goldman–Fristoe 2 Test of Articulation* from a standard score of 82 to 110. Finally, Participant I emitted 78 full echoics during the post–experimental probe condition and increased scores on the *Goldman–Fristoe 2 Test of Articulation* from 67 at pre–test to 103 in the post–assessment. Overall, participants required 180–220 trials to acquire a generalized auditory word match–to–sample and increased full echoic responses from an average of zero at pre–assessment to 75 in the post–experimental probes. Age equivalency scores on the standardized test of articulation improved an average of 2.5 years and standard scores increased an average of 33.75 points.

Speckman–Collins et al. (2007) evaluated the effects of the same auditory matching procedure used by Brown (2005) on listener, tact, and echoic responses in two young children with autism spectrum disorder. Participant A was 3 years, 5 months of age, and while she was able to match some items and objects following direct teaching, she did not exhibit a generalized matching repertoire. Similarly, she had been observed to produce true words but did not have a generalized echoic repertoire. Motor imitation was inconsistent and mands were limited to simple pointing and gesturing for preferred items. Language assessments indicated age equivalency scores of 1 year, 3 months for auditory comprehension and 11 months for expressive communication. Participant B was 4 years, 8 months of age. Vocal communication included mands, tacts, and autoclitics, although she did not exhibit spontaneous language across environments. She had a generalized echoic repertoire, but articulation errors were noted in her speech.

The auditory matching procedures consisted of five phases including matching of environmental sounds in the presence of no sounds, matching environmental sounds in the presence of sounds, matching words in the presence of nonsense words, matching words in the presence of words, and matching of words with similar phonetic structures. Hence, increasingly finer auditory discriminations were required as participants moved through the five phases of treatment. Four sets of picture stimuli were used to probe for listener, tact, and echoic responses in 20 trial sessions for each operant. Echoic probes were conducted after each phase of treatment to determine if the auditory matching procedures improved echoic behavior. Participant A did not produce any echoics in baseline. After phase two, she began emitting echoic responses. The number of echoics correctly produced in the final probe sessions ranged from 2–13 across stimulus sets. Accuracy of echoics showed less significant improvements across probe sessions for Participant B. In baseline, she produced five echoics. The number of correct echoics in the final probe sessions ranged from 6–20.

Choi et al. (2015) investigated the effects of an auditory match to sample procedure on echoic behavior in four elementary school students, aged 7–8 years old, with autism spectrum disorder. All children were described as having inexact echoic repertoires which were characterized by poor clarity in articulation or lack of point–to–point correspondence. The auditory matching task was presented on a computer. Three red circles were displayed on the screen, one at the top center, and two below, side–by–side. When activated via touch or mouse click, the top circle presented the sample spoken word or phrase. The two circles below presented either a matching or nonmatching spoken word or phrase. The matching tasks were arranged in a hierarchy of increasing difficulty. Training began with matching speech sounds of single words in the presence of similar words (e.g., plate vs. eight, walking vs. talking, carrot vs. caring). The next phases required matching phrases (e.g., red hat vs. blue hat, a fast dog vs. the fast dog, the book is on the table vs. the pen is on the table). As training progressed the distinction between the positive and negative exemplars became finer. The therapist touched the top circle and each of the bottom circles. The top circle was activated again, and the therapist instructed to participant to “match”. Praise and a preferred edible or token were presented for correct matches. Following an incorrect match, the therapist modeled the correct match.

Echoic accuracy was evaluated across 100 common English words and 40 unfamiliar Korean words (Choi et al., 2015). Because of their unfamiliarity, the Korean words would serve as a more rigorous test of correspondence between the modeled word and the participant’s production. During pre–intervention probes of the English words, participants correctly imitated an average of 68 words (range 51 – 80) out of the 100 words presented. After training, participants echoed an average of 89 words (range 72 – 98). Improvements in imitation of Korean words were observed as well. During pre–intervention probes, participants correctly imitated an average of 15 words (range 13 – 16) out of the 40 words presented. Following training, participants correctly imitated an average of 30 words (range 24 – 34).

In a second experiment, Choi et al., (2015), evaluated the auditory matching tasks on echoic behavior in three children, ages 7–8 years old, two of which had a diagnosis of autism. Because of the ceiling effect observed with the English words in Experiment 1, the authors created a new list of 50 English words. This new list consisted of 2–5 syllable words that were likely unfamiliar to the participants (e.g., welfare, kyanite, zincography, methodology). Changes in Experiment 2 also included additional phases of the auditory match–to–sample. In these phases the participants matched uncommon, multisyllabic words (e.g., biological, occurrence, randomize, thermoplastic). During pre–intervention probes of the English words, participants with autism correctly imitated an average of 25 words (range 18 – 33). After training, participants echoed an average of 45 words (range 42 – 49). Imitation of Korean words increased from an average of 16 words (range 15–17) before intervention to 35 words (range 33–37) after the auditory match–to–sample tasks.

Du et al. (2017) used procedures similar to those of Choi et al. (2015) to evaluate the effects of an auditory match to sample procedure on echoic behavior in preschool children. Three children participated in the study, two of which were on the autism spectrum. Both of these children demonstrated early speaker abilities. Specifically, they were able to emit single word mands and tacts following echoic prompts. The auditory matching protocol developed by Choi et al. (2015) was presented in an iPad app developed by the authors which offered greater accessibility for young children and those with fine motor challenges who were unable to use a computer mouse.

The auditory matching protocol was divided into 22 phases requiring progressively finer discriminations. The first phase taught the point–to match topography. Three cartoon pictures were presented with one at the top of the screen and two side–by–side at the bottom. One of the cartoon pictures on the bottom matched the picture at the top in both appearance and sound (e.g., two pictures of a dog and one picture of a car). In all other phases, the presentation was the same, but the pictures were all identical and unrelated to the auditory target (e.g., stars, smiley faces). After mastering the point–to match topography, students learned to match sounds and words against the absence of sound, white noise, and other sounds and words. Later phases included longer and more complex words, phrases, and sentences. As in the study by Choi et al. (2015), the teacher pressed the top button and the two buttons on the bottom. The top button was then pressed a second time and participants were instructed to "match". The app collected data. Correct responses were automatically reinforced with a fireworks animation and sounds of applause and cheering. Following an incorrect response, the app provided a prompt to "try again," and highlighted the matching buttons. The teacher presented the task again to provide participants an opportunity to respond independently. A partial physical prompt was provided on the third attempt.

Pre and post intervention echoic probes were conducted with the same list of 100 common English words used in Choi et al.'s (2015) first experiment (Du et al., 2017). During pre–intervention probes, Participant A emitted 1 full echoic and 1 partial echoic. Following the auditory match–to–sample procedures, Participant A produced 51 full echoics and 30 partial echoics. Participant B emitted eight full echoics and six partial echoics in the pre–intervention probes. During post–intervention probes, Participant B produced 81 full echoics and 18 partial echoics.

Acquiring a generalized auditory matching repertoire improved both full and partial echoic responses in most participants with limited vocal imitation (Brown, 2005; Choi et al., 2015; Du et al., 2017; Speckman–Collins et al., 2007). One possible reason for the observed effects of the auditory discrimination procedures is an unintended stimulus–stimulus pairing (Choi et al., 2015; Du et al., 2017; Speckman–Collins et al., 2007). Through the auditory match–to–sample tasks, the adult voices may have been conditioned as reinforcers which increased the reinforcing effect of listening to speech. As the participants became more attentive to the adult voices, and their responses came under phonemic stimulus control, they learned to accurately respond to the speech of others. Therefore, auditory match–to–sample may be a behavioral cusp that facilitates the acquisition of more advanced verbal behavior (Choi et al., 2015).

Vocal Imitation Training

Perhaps the most common treatment used to establish and improve echoic behaviors and speech sound disorders is vocal imitation training. Very simply, this procedure consists of a vocal model and differential reinforcement of approximated or precisely imitated responses (Carroll & Klatt, 2008; Civindi–Motta et al., 2017; Lovaas et al., 1966). The adult presents a target word or sound and asks the child to imitate (e.g., “Say, /u/”), and a preferred stimulus is delivered contingent on imitation of the model, or an approximation thereof. Although this describes the more conventional application, it may be modified to include additional procedures such as shaping, chaining, or prompting.

Dyer et al. (1987) followed developmental models of speech acquisition to improve articulation in a 13–year–old girl with autism using vocal imitation training. The *Goldman–Fristoe Test of Articulation* was administered, but because the participant was unable to independently tact the pictures, the therapist modeled the target word and recorded the child’s imitation. Based on the results, the following phonemic targets were selected for treatment: /p, d, m, b, g, v, t, ɵ/. All sounds were targeted in the final position of words. Sounds were paired so that, in each session, an early sound and a later developing sound were taught. Ten trials of each sound were presented at each session. The therapist modeled the word containing the target sound, and contingent on correct imitation, the therapist provided verbal praise and an unspecified primary reinforcer. In all pairings, the earlier developing sound met the mastery criterion (i.e., 80% correct across three sessions) in fewer trials than the later developing sound. For example, /t/ which typically emerges between 2–6 years of age, was mastered in 40 trials while /ɵ/ which is usually acquired between 4.5 and 7 years of age, was mastered in 90 trials. Over time, fewer trials were needed to achieve mastery, suggesting generalization of learning. In conclusion, developmental norms should be considered when selecting targets for speech interventions.

Koegel et al. (1988) used a reversal design to compare the effects of two different reinforcement conditions, verbal attempts and motor speech, on vocal imitation. Four children, age 3–11 years, participated in the study. Three children were diagnosed with autism spectrum disorder and one child was classified as developmentally delayed with autistic characteristics. Standardized assessments were not administered, but all the children were described as nonverbal with deficits across multiple developmental areas. Current level of functioning for all participants was estimated to be at the 1–year level. Target vocalizations were selected based on their contextual fit with a variety of toys. One and 2–syllable words were selected for Participants A and D (e.g., “hit,” “music”), and simple phrases (e.g., “I want it,” “more blow”) were selected for Participants B and D who were vocally more advanced. At the start of each session a variety of toys, games, and activities were made available, and the participant was encouraged to select an item. The instructor interacted with the selected item and modeled the target vocalization. If the participant emitted the target word, or approximated it depending on the condition, the instructor delivered the selected item.

In the verbal attempts condition, any vocal response, regardless of its phonetic topography or similarity with the modeled vocalization, was reinforced with access to the selected item (Koegel et al., 1988). Conversely, in the motor speech or vocal imitation condition, only successive approximations of the target vocalization were reinforced. At the start of each session, the instructor conducted five echoic trials of each target vocalization. The child’s most frequent response served as the phonetic criterion for that session. When the child met or exceeded the criterion for six out of 10 responses, the criterion changed to require at least one more distinctive feature (i.e., the articulatory or acoustic elements of the phonemes). Each condition lasted two to six sessions before the child entered the other condition. The children participated in each condition 3–6 times for a total of 14–24 sessions.

Vocalizations were scored on a scale of 0–12 based on the distinctive features included in the child's response (Koegel et al., 1988). The average score for each target word was calculated at each session. All participants showed the greatest improvements in the verbal attempts condition. More specifically, all the children always showed improvements in the verbal attempts condition while in the motor speech condition, performance varied with participants demonstrating smaller gains, no changes, or deterioration in speech production.

Esch et al. (2005) used a multiple baseline across behaviors to evaluate the effects of stimulus–stimulus pairing and direct reinforcement on echoic responses. Participants ranged in age from 6 years, 10 months to 8 years, 2 months and all carried a diagnosis of autism spectrum disorder. Based on assessments of speech and language, all participants had receptive and expressive language ages below 12 months and did not exhibit imitation of any vocalizations. Free–operant vocal behaviors included vowels and, for one participant, occasional production of consonant–vowel syllables. Following observation, target responses were selected based on these spontaneous productions and approximations of sounds. During the stimulus–stimulus pairing, the target sound was presented every second for 3 s followed by delivery of a preferred stimulus. Immediately following 30 pairing trials, 10 echoic probes, or vocal imitation trials, were conducted. The target sound was modeled, and a preferred stimulus was delivered contingent on imitation or approximation within 5 s. Overall, these procedures were unsuccessful. None of the participants showed an increase in echoic responses following the pairing procedures for any of their three target sounds.

Across two experiments, Carroll and Klatt (2008) also evaluated the effects of stimulus–stimulus pairing procedures and differential reinforcement on echoic behavior. Two children, age 23–24 months, with a diagnosis of autism spectrum disorder, participated in the study. The *Behavioral Language Assessment Form* was used to identify current communication levels for both participants. Participant A did not exhibit mands or tacts, and while she emitted a few sounds spontaneously, none were under echoic control. Participant B had a strong echoic and imitative repertoire, was able to mand for preferred items, and could tact simple objects. Following repeated observations of vocalizations emitted during therapy sessions, both a known target and a novel target were selected. The known target was a one–syllable vocalization produced at the lowest frequency during observation. The novel target was a vocalization never produced by the participant during observation. In Experiment 1, a multiple baseline across behaviors was used to evaluate the effect of stimulus–stimulus pairing on frequency of vocalizations. A total of 20 pairing trials were conducted each session. The experimenter presented the target sound three times, delivered a preferred stimulus, and presented the target sounds two additional times. After the participant engaged with the preferred stimulus for 20 s, the next pairing trial was performed. The frequency of target sound productions and approximations were recorded during 5–min observation sessions both before and after the stimulus–stimulus pairing trials.

For Participant A, increases in production of the known target sound “ts” occurred in both the presession and postsession observation periods as compared to baseline (Carroll & Klatt, 2008). These increased rates were maintained across subsequent observation periods in the absence of the stimulus–stimulus pairing. However, the novel sound, “buh,” was never emitted in the baseline, pairing, or maintenance condition. Because the stimulus–stimulus pairings did not produce an increase for either of the target vocalizations for Participant B, a direct reinforcement procedure, or vocal imitation training, was implemented for the known vocal target. A total of 20 echoic trials were presented each session whereby the experimenter modeled the known target sound and delivered a preferred stimulus contingent on production of the sound within 3 s. While direct reinforcement yielded imitation with 75% – 100% accuracy, an increased frequency of the target sounds did not occur in pre– and post– session observation periods.

The stimulus–stimulus pairing produced an increase in the known target vocalization for Participant A (Carroll & Klatt, 2008). Therefore, the purpose of Experiment 2 was to bring the target vocalization under echoic control through direct reinforcement. A comparative design was used to evaluate five variations of direct reinforcement procedures. The final two variations yielded the highest percentage of echoics. In the fourth procedure, all vocalizations were reinforced, but imitation of the target vocalization received more immediate reinforcement. A total of 15 echoic trials were dispersed across one hour treatment sessions. A preferred stimulus was immediately delivered following correct imitation or approximation of the target vocalization. If any other vocalization followed the experimenter’s model, a preferred stimulus was delivered following a 5 s delay. In the fifth procedure, a preferred stimulus was delivered contingent on imitation of the target vocalization. No other vocal behaviors were reinforced. With these two variations, the target vocalization was imitated an average of 38.5% and 81% of trials, respectively. Imitation was maintained at 100% correct across all trials for 10 weeks following treatment. Anecdotally, the authors reported acquisition of 12 additional echoic responses following mastery of the known vocal target in the study.

Chaining. Tarbox et al. (2009) used a modified chaining procedure within vocal imitation training to increase the complexity of existing echoic behaviors. The study included three participants, two of which had a diagnosis of autism. Participant A was 5 years old. He independently used one syllable, vocal mands, and with prompts, could produce one syllable vocal tacts. Participant B was 7 years old. He had an echoic repertoire of approximately six sounds, but did not exhibit any independent, functional vocal communication. Three vocal targets were selected for each participant in accordance with their current therapeutic needs, and a multiple baseline across target behaviors was used to evaluate changes as a result of the treatment. To implement the chaining procedure, targets were divided into two components (e.g., “Monday” was divided into “mon” and “day”). The first component was modeled on the first trial (e.g., “mon”), the second component on the second trial (e.g., “day”), and the entire vocal target was presented on the third trial. A total of 15 trials were presented each session. A preferred stimulus was delivered following correct imitation of the model on each trial. Following an incorrect response, the trial was repeated once. Both participants improved production of the full target echoic, and most were maintained after treatment was withdrawn. The results indicate this modified chaining procedure may effectively increase the length and complexity of echoic responses. Interestingly, Participants A met the mastery criterion for each vocal target in nine or fewer sessions. Participant B required 35 treatment sessions to master the first vocal target and correctness was not maintained following treatment withdrawal. However, the second and third vocal targets were mastered in less time; 17 and six treatment sessions, respectively. After mastery of these subsequent targets, maintenance of the first echoic target improved and eventually reached 100% correct.

Mand Model. Drash et al. (1999) used a mand–model procedure to teach echoic behavior. Participants included three children on the autism spectrum, ages 2 years, 6 months to 3 years, 6 months. All the children were described as nonverbal and, based on language assessments, had age equivalency scores between 10–16 months. Participant A spontaneously emitted several sounds and words but had no functional use of language and did not imitate vocal behavior. Participant B imitated two sounds and three words with 25% accuracy and inconsistently responded to prompts for vocal mands (53% of trials). Participant C imitated four sounds and five words with 54% accuracy and responded to prompts for vocal mands on 95% of trials. Speech–like vocal behaviors under control of a motivating operation were differentially reinforced. That is, crying or screaming were never reinforced, but when the child reached for a desired item, the therapist asked a question such as, “Do you want this?” and any appropriate vocalization that followed was reinforced.

After teaching vocal mands, the mands were used to teach echoic behaviors (Drash et al., 1999). The authors identified speech sounds that occurred at a high frequency (e.g., /ɑ/, /bʌ/, /i/, /m/) and then selected reinforcers with an associated vocal topography that contained these sounds. When the child emitted “mmm,” he was given a chocolate candy, and when he vocalized “ah,” he was given a bite of an apple. The therapist imitated the child’s vocalization and before delivering the reinforcer, prompted him to imitate. This way, the vocal response came under control of the therapist’s vocal behavior as well. After seven training sessions, the children acquired echoic repertoires of 4+ sounds and 9+ words imitated with 70% accuracy.

Civindi–Motta et al. (2017), using an adapted alternating–treatments design, compared three different teaching procedures for echoic behavior and conducted functional analysis probes to identify the function of vocalizations. The specific teaching procedures were vocal imitation training, stimulus–stimulus pairing, and a mand–model procedure. In vocal imitation training, a sound is modeled, and a preferred stimulus is delivered contingent on correct imitation of the model or on successive approximations of the sound. Similarly, in the mand–model procedure, a preferred item is delivered following a correct verbal response. The preferred item is made available and as the child attempts to access it, the therapist mands for the desired response (e.g., "What do you want?"). The therapist may also model the desired vocal response (e.g., "cookie"). In the stimulus–stimulus pairing procedure, the target sound is paired with a conditioned or unconditioned reinforcer to increase spontaneous production. Differential reinforcement is then used to bring the target sound under echoic control.

The study included six participants with various developmental disabilities ranging from 7 – 17 years of age (Civindi–Motta et al., 2017). Preassessment echoic behavior was described as absent, limited, or poor correspondence for each participant. Target sounds were selected following echoic assessment. When possible, sounds belonging to the same phonemic category (e.g., long vowels) with similar ages of acquisition in typical development, were assigned to each of the three teaching procedures. According to the results, the efficacy of each teaching procedure varied across participants. Participants A, C, and D had an echoic repertoire described as "limited" or "poor correspondence". Vocal imitation training produced consistently higher levels of responding for each of these three participants, and the functional analysis confirmed that these sounds were under echoic control. While the mand–model procedure also effectively increased echoic responses for these three participants, it did not produce the level of responding observed with the vocal imitation training. For Participant B, described as having "limited" echoic behavior, only the sounds taught with the mand–model procedure were acquired. Participant D did not exhibit echoic behavior prior to the study. Both the vocal imitation training and stimulus–stimulus pairing procedures resulted in increased responding, although levels were higher under stimulus–stimulus pairing. Interestingly, these results were not replicated in a second comparison, where Participant D failed to acquire any echoic responses under any of the teaching procedures. Participant E had limited echoic behavior prior to the study. He rarely emitted the target sounds under any of the teaching conditions, leaving the authors to conclude that none of the procedures were effective in teaching him echoic behavior.

Physical Prompts

Lovaas et al., (1966) described a discrimination procedure for improving vocal imitation. The procedure was used with two 6–year–old children on the autism spectrum, who emitted occasional vowel sounds with no clear communicative intent. First, any vocal behavior was reinforced, as was looking directly at the adult's mouth. Secondly, vocalizations that followed the adult's vocal stimulus within 6 s were reinforced. In the third phase of treatment, the authors selected a vocal target that could be physically prompted. For example, when teaching /b/, the adult would model the sound and simultaneously hold the child's lips closed and quickly release when the child was observed to exhale. The prompt was faded by reducing contact with the child's lips until the child responded independently following the adult's vocal model. New sounds were introduced in the fourth phase of treatment, and, to ensure the responses were under control of the antecedent vocal stimulus, previous sounds were randomly presented. Data for one of the participants showed that he progressed from imitating a single sound in isolation on day 10 and over the next 16 days acquired several syllables and words such as "boy," "ball," "why," “no,” and “bread.” The authors noted that, after 26 days of treatment, both boys were imitating words with “such ease and rapidity,” that treatment focused on using the newly acquired vocalizations in a linguistically functional manner (p. 706).

Aravamudhan and Awasthi (2020) used sufficient–response exemplar training to improve speech productions of two children on the autism spectrum, ages 11 and 15 years. Participant A could articulate most words but had difficulty with certain consonant blends in the initial position of words. Participant A had a functional speech repertoire of 30–word approximations that were only understandable to familiar listeners. She was described as having a profound speech sound disorder. Consonant blends, /st/, /sp/, /sm/ and 10 corresponding word sets for each, were selected for Participant A. Consonant blends, /st/, /sp/, /sc/ and 10 corresponding word sets for each, were selected for Participant B. Treatment procedures included within–stimulus prompts, lip–tongue–teeth position prompt, and chaining procedures. The within–stimulus prompt was an exaggerated model of the part of the word the participant had difficulty imitating. For example, when presenting the word “stop” the therapist would say, “ssssstop”. The lip–tongue–teeth position prompt was implemented to provide a model of the appropriate place and articulatory movement. For example, the therapist modeled the labiodental position for producing the target /ʌv/ and tongue retraction for the target sound /k/. A spoon was used to gently facilitate retraction of the tongue. Implementing the chaining procedure required that the target words be divided into two parts (e.g., the word spoon was split into “spoo” and “n”). The first part was modeled, the participant imitated, the second part was modeled, and the participant imitated. The time between presentations was gradually reduced until the full word was modeled and imitated. The therapist presented the target word (e.g., Say “stop”) and provided the predetermined prompt. If the participant correctly imitated the target word, a token was provided. Following an incorrect response, the target word was presented again. If the participant correctly responded, a token and social praise were provided. When an error occurred on the second presentation, the therapist asked the participant to imitate a mastered vocalization and provided a token for a correct response.

Participant A took on average, 532 trials (range 72–1230), within 54.75 sessions (range 6–125), over 9.75 days (range 1–23) to master the target words in training (Aravamudhan & Awasthi, 2020). After mastering the first word in the /st/ and /sp/ set, the correct articulation generalized to other, untrained words in the set and the same generalization occurred in the /sm/ set after mastering the first two targets. Participant B took on average, 730 trials (range 264–2044), within 77.11 sessions (range 30–201), over 13.88 days (range 5–36) to master the target words in training. Training was provided on the first six words of the /st/ set and correct articulation generalized to two untrained words. Three words from the /sp/ set were trained and correct articulation generalized to five untrained words in the set. Due to repeated illness the /sc/ set was not trained. Although both participants improved articulation above baseline measures, and their articulation generalized to untrained words, Participant B required more trials and more training words to achieve mastery, and in many cases was only able to achieve a close approximation of the target word. The authors explained that these differences may be due to the severity of her articulation impairment. Thus, additional prompts and simpler approximations may reduce instructional time for children with severe speech sound disorders.

Using the principles of precision teaching with lip–tongue–teeth position prompts, Aravamudhan and Awasthi (2021) taught Participant B from the study by Aravamudhan and Awasthi (2020), to imitate /fʌ/, /kʌ/, and /θʌ/. With the goal to establish fluent echoics with point–to–point correspondence, each session lasted only 15–30 s. For the first three days of intervention, the lip–tongue–teeth prompts were provided in the first three trials to teach the correct placement of articulators for each syllable. The prompts were reinstated after two consecutive errors for an additional three trials. A physical prompt was used to facilitate tongue retraction for /kʌ/. The participant was taught to touch her index finger to her tongue tip and gently push back approximately 1 millimeter. Feedback (e.g., "yes") was also provided for correct responses on a variable ratio schedule of two. At the fourth day of intervention, the prompts were only applied after two consecutive errors, and feedback was not provided. Access to a preferred stimulus was provided at the end of each session. If the rate of incorrect responses did not decrease over time, an accuracy–building phase was added, whereby prompts were reinstated after each error. Because the target syllable /kʌ/ was frequently substituted with /θʌ/, a priming phase was added. Prior to each session the therapist presented the target, the predetermined prompt, and feedback (e.g., "correct" or "try again"). The session began after three correct responses.

Participant B achieved the mastery criterion of 60 correct echoics per minute with fewer than five errors (Aravamudhan & Awasthi, 2021). The target syllable /fʌ/ was mastered in only 42 min of training across 18 days, /kʌ/ was mastered in 24 min of training across 16 days, and /θʌ/ was acquired in 30 min across 10 days. Training time was substantially reduced as compared to Aravamudhan and Awasthi (2020). Thus, these authors conclude that these procedures are efficient and effective for improving echoic behavior in an adolescent on the autism spectrum with a severe speech sound disorder.

PROMPT. *Prompts for Restructuring Oral Muscular Phonetic Targets* (PROMPT) is a multidimensional approach to the assessment and treatment of speech sound disorders (Hayden, 2004; L. Slim, personal communication, 2022). Perhaps its most salient feature is the application of physical prompts to the muscles and structures necessary for articulation including the cheeks, lips, and chin (Chumpelik, 1984; Dale & Hayden, 2013). With a separate prompt for each English phoneme, the prompts provide information on place of contact, jaw height, nasality, voicing, as well as transitions, stress, timing, and sequencing of sounds and syllables. Prompts may be used at the sound, syllable, word, or phrase level. In this way the child learns speech sound production without having to fully control the motor movements.

Four levels of prompts comprise the PROMPT approach: parameter prompts, syllable prompts, complex prompts, and surface prompts (Hayden, 2004, 2021; L. Slim, personal communication, 2022). Parameter prompts provide support and stability to teach dissociation of movements (Hayden, 2004, 2021). Speech motor actions, such as phonation, mandibular range, and labial rounding/retraction, are targeted. Syllable prompts shape production of CV and VC syllables. The mandible is positioned and supported according to the vowel sound, so that the speaker can access labial and lingual movement. Complex prompts are often used to teach sounds in isolation. They provide information about multiple aspects of sound production including jaw height, labial–facial movements, lingual position, and timing. Surface prompts provide the most critical, but least amount of information, as compared to the other prompt levels. They are useful for teaching co–articulatory transitions and coordinated movements for speech production.

Broader than just physical prompts though, PROMPT is a holistic model with a conceptual framework that recognizes the interaction of multiple internal and external variables influencing speech production, including cognitive–linguistic skills and experiences, physical–sensory structures, and social–emotional acts (Hayden, 2004, 2006, 2u021; L. Slim, personal communication, 2022). Consequently, deficits in any one of these errors can disrupt speech production. To assess and treat speech sound disorders, PROMPT follows a model of motor speech function called *The Motor Speech Hierarchy* (Hayden, 2004, 2006, 2021; L. Slim, personal communication, 2022). As the name implies, it depicts a hierarchical arrangement of seven levels of motor speech development and control. The Motor Speech Hierarchy begins with Stage 1, *Muscle Tone*, which considers stability and alignment of the head, neck, and trunk because postural control is necessary for speech production. Stage 2, *Phonatory Control*, requires that the individual initiate and sustain voicing while the jaw is moving. Stage 3, *Mandibular Control or Vertical Plane of Movement*, examines jaw movement, stability, and symmetry. This is crucial for dissociation and grading of articulatory movements. Stage 4, *Labio–Facial Control or Horizontal Plane of Movement*, considers dissociation of the lips and jaw. That is, the lips should move independently of the jaw. Stage 5, *Lingual Control or Anterior–Posterior Plane of Movement*, involves dissociation of the tongue and jaw. Stage 6, *Sequenced Movement*, evaluates sound sequences across planes of movement. Stage 7, *Prosody*, is the final stage of the Motor Speech Hierarchy, which addresses rate, stress, duration, and other suprasegmental features of speech. To treat speech sound disorders, the lowest level impacted must be addressed first because subsequent stages are directly affected by preceding levels.

PROMPT is described in three treatment phases (Hayden, 2004). Phase 1 is for individuals with limited speech. Here, treatment teaches motor control in a single plane of movement according to Stages 2, 3, and 4 of the Motor Speech Hierarchy. Phonation, jaw gradation, and lip rounding/retraction are addressed through production of bilabial consonants and vowel sounds with varying degrees of jaw height. In Phase 2, multiple levels of prompts are used to teach sound sequencing across two planes of movement. Stages 3, 4, and 5 of the Motor Speech Hierarchy are addressed as the speaker learns more refined articulatory movements and dissociation of jaw, lips, and tongue. Finally, phase 3 addresses Stages 5, 6, and 7 of the Motor Speech Hierarchy. Articulatory patterns are further refined by teaching movement across three planes and focusing on increasingly complex syllable structures and co–articulation. Surface prompts are common in this phase, as are "air prompts," where the clinician provides the prompt near the facial structures, but without direct contact (Hayden, 2004; L. Slim, personal communication, 2022).

PROMPT has been evaluated in a variety of populations with evidence for improved speech production in children with severe speech sound disorders (Namasivayam et al., 2013; Yu et al., 2014; Yu et al., 2018), motor speech impairments (Dale & Hayden, 2013; Namasivayam et al., 2020), cerebral palsy (Ward et al., 2013; Ward et al., 2014), bilateral cleft lip and palate (Herreras Mercado et al., 2019), and some preliminary evidence exists for adults with aphasia and apraxia of speech (Bose et al., 2001; Freed et al., 1997). To date, only one study has investigated the effects of PROMPT in children on the autism spectrum (see Rogers et al., 2006).

Using an AB experimental design, Rogers et al. (2006) evaluated the effects of PROMPT and the Denver Model on vocal communication in nonverbal (i.e., fewer than five spoken words per day) children on the autism spectrum. The study included 10 children, ages 20–65 months, who were randomly assigned to treatment conditions. They received either PROMPT or Denver Model for 1 hour per week for a total of 12 weeks. Parents were also asked to incorporate treatment objectives into daily routines. Frequency of approximations, words, and phrases were measured in 10–15 min play sessions across baseline, treatment, and follow–up. By the end of treatment, 8 out of 10 participants acquired at least five functional, spontaneous words. Children receiving PROMPT gained an average of four raw score points (range 1–13) on the expressive language subtest of the *Mullen Scales of Early Learning* after treatment and those in the Denver

Model condition gained an average of 3.4 raw score points (range 0–10). Parents reported the number of words produced and understood before and after treatment. The children in the PROMPT condition reportedly increased from an average of 15.6 words (range 0–40) before treatment to an average of 69.6 words (range 2–190) following treatment. The children in the Denver Model condition reportedly increased from an average of 7.4 words (range 0–20) before treatment to an average of 50.4 words (range 4–193) after treatment. Although the results suggest that PROMPT may be a promising treatment for children on the autism spectrum, the inherent risk of type 1 error in the AB design prevents conclusions regarding the effects on vocal communication.

Clinically, vocal imitation training seems to be the most common approach to establishing and/or improving speech production in children on the autism spectrum, yet the results are mixed (e.g., Civindi–Motta et al., 2017). Some studies have identified new vocal behaviors following treatment (e.g., Dyer et al., 1987), especially with additional modifications (e.g., Tarbox et al., 2009), while other investigations have been unsuccessful (e.g., Esch et al., 2005). Given the importance of vocal imitation in the development of social communication and expressive language (Aravamudhan & Awasthi, 2020; Pecukonis et al., 2019), the number of children that exhibit minimal vocalizations even after intervention (Norrelgen et al., 2015), and the social significance of vocal communication (Mayo et al., 2013), it is essential that we identify effective techniques for teaching vocal–verbal behavior to children on the autism spectrum.

Purpose of the Current Study

Physical prompts as a supplementary stimulus in vocal imitation training has shown positive gains in speech production for children on the autism spectrum (e.g., Lovaas et al., 1966; Aravamudhan & Awasthi, 2020; Aravamudhan & Awasthi, 2021), although it remains a relatively understudied method. Use of procedures such as these may address underlying motor planning and coordination challenges believed to cause many of the vocal communication deficits associated with autism (see Beiting, 2022; Chenausky et al., 2019; Dawson, 2000). Therefore, the purpose of the current study was to evaluate the effects of physical prompts on echoic behavior in children on the autism spectrum. More specifically, conventional vocal imitation methods were compared to treatment that included the use of physical prompts, based on the PROMPT method, in both children with reduced speech intelligibility and those with minimal vocal behavior. The impact of both procedures was evaluated for the effect on vocal imitation.

Method

Participants

The study included four children, each with a diagnosis of autism spectrum disorder. To be eligible for participation, all children had to imitate non–vocal motor actions, engage in noncontextual, spontaneous vocal behavior, and exhibit severe deficits in echoics as demonstrated by standardized assessment. All participants carried a diagnosis of autism spectrum disorder and attended a school for children with autism and other developmental disabilities, where they received interventions based on the application of behavior analytic principles and other therapeutic services.

Sarah was a White, 10–year–old female who received 35 hours a week of interventions based on applied behavior analysis and 30 min each of speech–language therapy and occupational therapy. She reportedly imitated a variety of non–vocal motor actions and simple vocalizations. During play, Sarah engaged in spontaneous vocal behaviors including humming, singing, and approximating phrases from her favorite television shows. Her mother reported hearing a variety of sounds in her vocal play including /h/, /ɑ/, /n/, /i/, /g/, /æ/, and /s/. Sarah's social communication was predominantly mands. At school and during therapy sessions, she used a speech–generating device to communicate with teachers, therapists, and peers. Although she was described as very independent at home, (e.g., rarely making requests, independently accessing desired items/activities) her mother explained that she learned to vocally approximate the "I want" carrier phrase to make requests.

Jayden was a Caucasian, 11–year–old male who received 30 hours of interventions based on applied behavior analysis, 30 min of occupational therapy, and 90 min of speech–language therapy each week. Due to staffing issues, Jayden's speech–language therapy was temporarily suspended and was later resumed during the course of the study. His speech–language therapy primarily focused on alternative/augmentative communication via a speech–generating device. Jayden imitated several non–vocal motor actions, and his mother explained that he would attempt any action he was asked to imitate. When asked what sounds she observed in spontaneous vocal play, Jayden's mother replied, "Lots of sounds"—squeals, vowels, and reduplicated babbling, such as /nənənə/, characterized his vocal behavior. Echoics included /ɑ/, and /m/ and /h/, though inconsistently. His social communication predominantly consisted of mands for items/activities and access to the restroom using a single icon or carrier phrases on his speech–generating device. Jayden also utilized conventional gestures, such as pointing and head shaking/nodding. These gestures were, at least in part, controlled by motivating operations. Vocally, his mother stated that he approximated "hi" and "yeah."

Benjamin was a White, 5–year–old male who received 30–35 hours of interventions based on applied behavior analysis therapy, 1 hour of occupational therapy, and 30 min of speech–language therapy each week and was to begin physical therapy soon. According to his mother, Benjamin was able to use signs, such as “eat,” “water,” “more,” “all done,” and “help,” although he often indicated his wants and needs through other means. For example, she remarked that Benjamin independently sat at the table when he wanted to eat. Parents were awaiting insurance authorization for a speech–generating device, which was later obtained, and his treatments focused on communication via the new device. Benjamin reportedly imitated simple motor actions such as clapping, touching his head, touching his nose, and pointing. His vocalizations consisted of squeals, shrieks, laughs, and “clicking noises.” Specific sounds included /b/, /d/ and /w/ in consonant–vowel (CV) syllables (e.g., “duh,” “wuh”), and /m/ in reduplicated syllables, such as “mumu.” Vocal imitation was described as weak, although his mother explained that he occasionally imitated /m/ and “clicking sounds.”

Theo was a Caucasian, 3 year, 3–month–old male who received interventions based on applied behavior analysis for 24–30 hours per week, 1 hour of physical therapy per week, and 30 min each of occupational therapy and speech–language therapy twice weekly. He vocally approximated 1–2–word utterances. His father reported a variety of sounds and words in vocal play and explained that Theo could imitate numerous motor actions and vocal sounds. He expressed concerns regarding Theo’s articulation. Specifically, he noted final consonant deletion, which interfered with his speech intelligibility and overall vocal communication.

Setting

All assessments, treatment, and probes were conducted at the participants’ school. The children were each seated in a chair appropriate to their height (i.e., feet planted on the floor with hips, knees, and feet forming an angle of approximately 90 degrees), and the therapist sat in a smaller chair or on the floor. This seating created a height differential which placed the child’s head in a slightly downward position, allowing the therapist to administer the physical prompts more easily (see Hayden et al., 2015). A table, or other flat surface was nearby to hold data collection sheets, preferred stimuli, and computer. The child and therapist sat on the same side of the table, in close proximity, facing each other. All sessions were video recorded on a personal computer for later reliability scoring.

Assessments

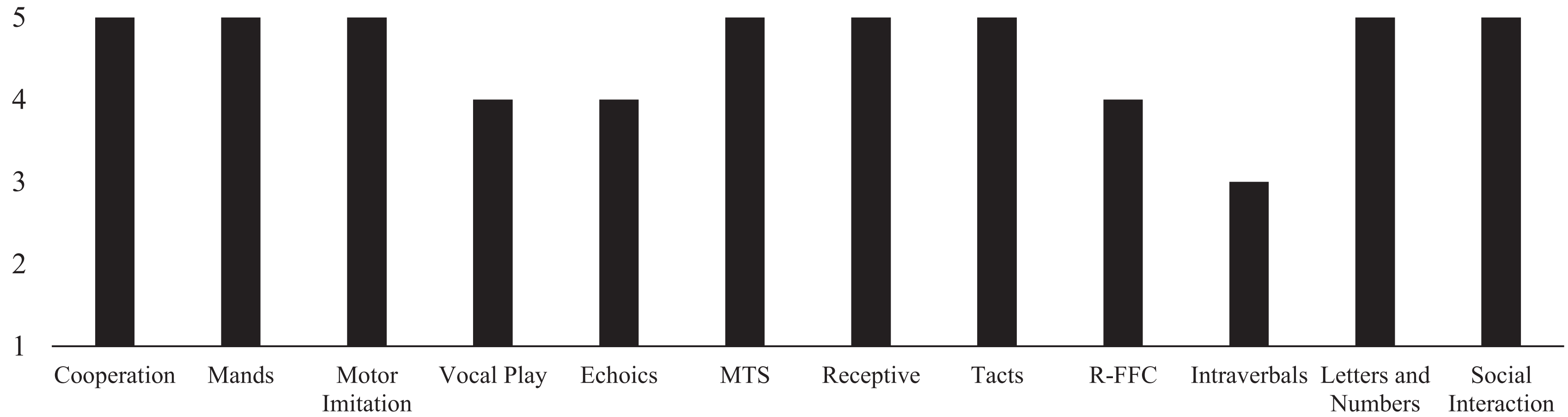
Behavioral Language Assessment

Each participant's supervising Board Certified Behavior Analyst (BCBA) completed the Behavioral Language Assessment Form (BLAF) to provide a brief overview of the child's verbal behavior (Sundberg & Partington, 1998; see Esch et al., 2005). The BLAF consists of 12 questions covering verbal operants and other language related skills. For each question, informants select one of five response options that best describe the child's typical level of performance. A score of one, for any given item denotes the absence of the behavior, and a score of five, indicates the behavior was observed at a relatively higher level. Accordingly, scores of 1–2 represent early learners, scores of 3–4 intermediate learners, and scores of 4–5 advanced learners.

Sarah received the highest score (5) for each of the following areas: cooperation with adults, motor imitation, matching–to–sample, receptive, mands, and tacts. Sarah's supervising BCBA indicated that she worked well at a table without disruptive behavior for 10 min although she was sometimes less cooperative with unfamiliar adults. She easily, and sometimes spontaneously, imitated fine and/or gross motor movements and could match most items and copy block designs. Receptively, Sarah could point to at least 100 items, actions, persons, or adjectives. When given four functions, features, or classes, she could identify 25 items. She could expressively label over 100 items or actions in the context of short sentences. Using 10 or more words, signs, or pictures, Sarah frequently requested preferred items and activities. Although she was not reported to regularly initiate and sustain verbal interactions with peers, Sarah was able to verbally interact with others when prompts were provided. Intraverbal behavior received the lowest score of all items on the BLAF. She could reportedly fill in 10 non–reinforcing phrases or answer at least 10 simple questions. Sarah's spontaneous vocal play was described as frequent with varied intonation and a few words, although not necessarily clear and understandable. She repeated or closely approximated many different words, although again, her vocal productions were not necessarily clear and understandable. Overall, for most areas of language, Sarah was classified as an advanced learner with scores of 4–5 on the BLAF (see Figure 1).

Figure 1

Results of Sarah's BLAF

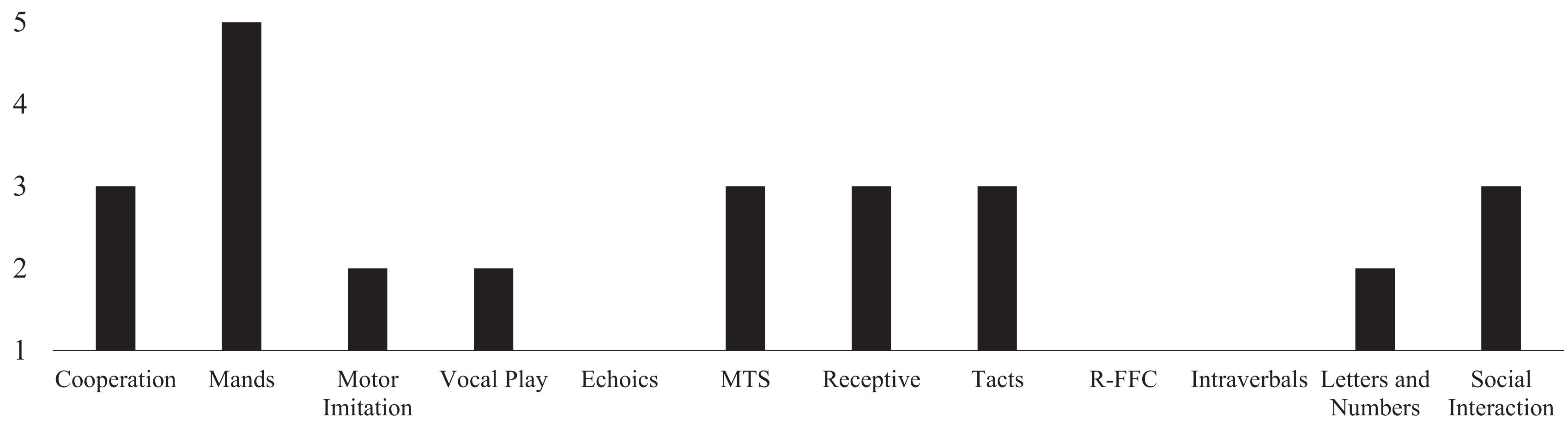


Note. Informant ratings across developmental domains based on a scale of 1–5. BLAF = Behavioral Language Assessment Form; MTS = Match to Sample; R–FFC = Receptive by Function, Feature, and Class

Jayden earned the highest score (5) for mands. He frequently requested preferred items and activities using 10 or more words, signs, or pictures. All other BLAF items earned scores from 1–3. His social interactions were primarily to ask adults for reinforcers. Jayden reportedly cooperated with adults by giving five responses without disruptive behavior. He reportedly imitated a few gross motor movements modeled by others. Receptively, Jayden followed a few instructions to complete actions or touch items and could identify at least three letters or numbers. However, he could not identify any items based on function, feature, or class. He could match 5–10 objects or pictures to a sample, and label 6–15 items or actions. Early intraverbal behavior, such as filling in missing words or parts of songs, was not reported. Jayden's spontaneous vocal play included a few sounds produced at a low rate. With a score of 1, his echoic behavior received one of the lowest scores on the assessment, indicating he could not repeat any sounds or words. Overall, Jayden's scores on the BLAF ranged from 1–3 (see Figure 2).

Figure 2

Results of Jayden's BLAF

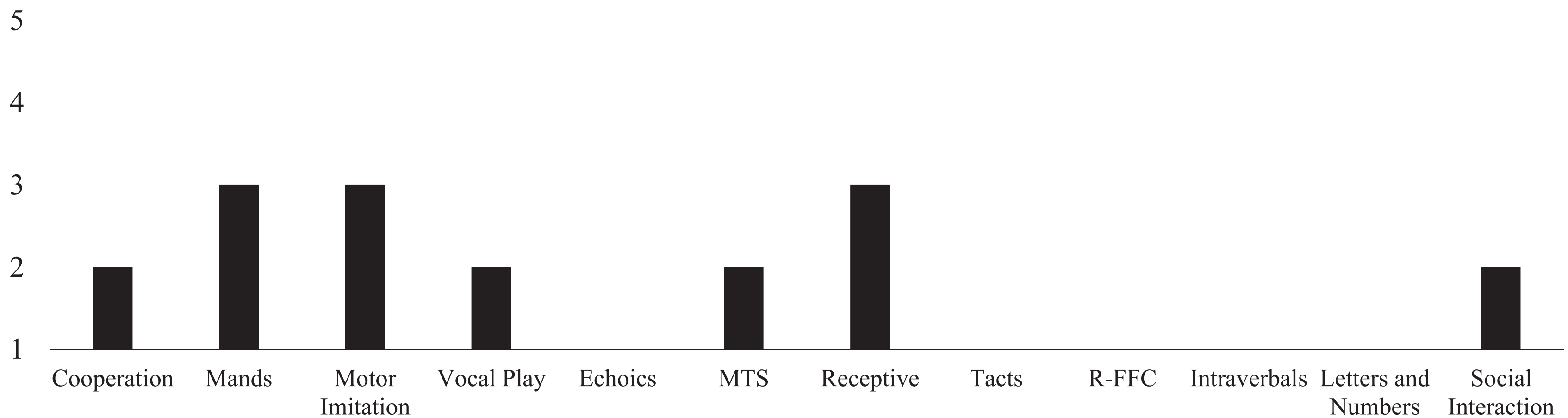


Note. Informant ratings across developmental domains based on a scale of 1–5. BLAF = Behavioral Language Assessment Form; MTS = Match to Sample; R–FFC = Receptive by Function, Feature, and Class

Benjamin’s functional expressive language was limited to mands. His supervising BCBA indicated that he used signs for “more” and “eat” to ask for reinforcers. He did not label items or actions, fill in missing words or parts of songs, or repeat any sounds or words. He initiated social interactions by physically approaching others. Receptively, he followed a few instructions to complete actions or touch items. He could not identify letters, numbers, written words, or items based on function, feature, or class. Benjamin could match 1–2 objects or pictures to a sample. Upon request, Benjamin imitated several gross motor movements. He was cooperative with adults but would typically complete only one brief and easy response for a powerful reinforcer. Benjamin reportedly produced a few speech sounds in vocal play, although at a low rate. For most areas of language, he was considered an early learner, earning scores of 1–2 on the BLAF (see Figure 3).

Figure 3

Results of Benjamin's BLAF

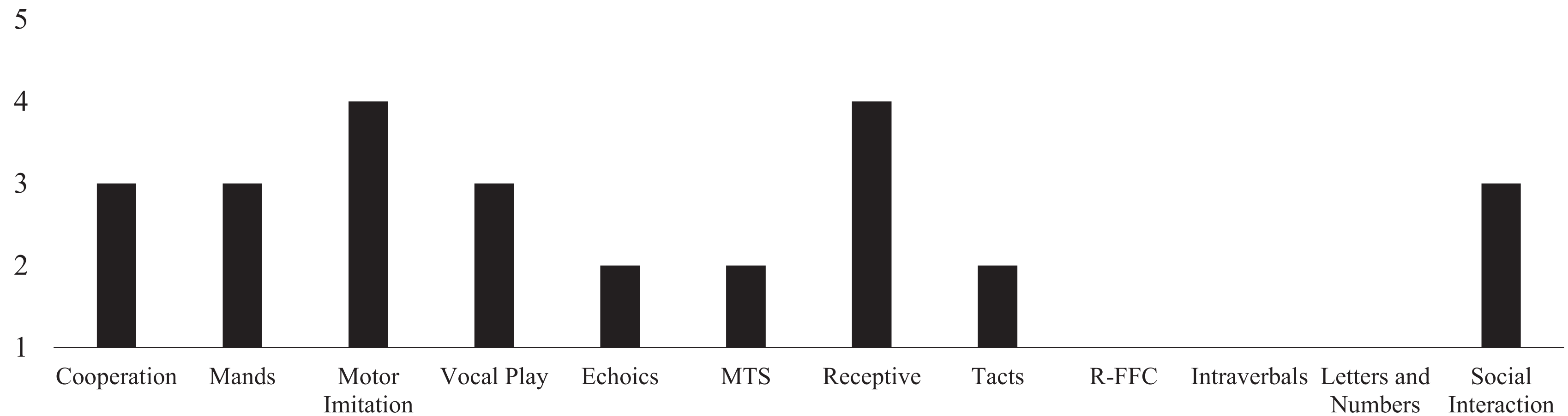


Note. Informant ratings across developmental domains based on a scale of 1–5. BLAF = Behavioral Language Assessment Form; MTS = Match to Sample; R–FFC = Receptive by Function, Feature, and Class

Theo's strongest areas of language and related skills were motor imitation and receptive language. His supervising BCBA indicated that Theo imitated several fine and gross motor movements on request, followed many instructions, and pointed to at least 25 items. Theo cooperated with adults and would typically give five responses without disruptive behavior. Using 1–5 single vocal words, he readily asked adults for reinforcers. He spontaneously vocalized many speech sounds with varying intonation and imitated a few specific sounds or words. Theo could match 1–2 objects or pictures to a sample, but could not yet identify letter, number, written words, or items based on function, feature, or class. He labeled 1–5 items or actions, but had not acquired intraverbal behavior. Theo presented as an early to intermediate learner on most language skills and related areas on the BLAF (see Figure 4).

Figure 4

Results of Theo's BLAF



Note. Informant ratings across developmental domains based on a scale of 1–5. BLAF = Behavioral Language Assessment Form; MTS = Match to Sample; R–FFC = Receptive by Function, Feature, and Class

Kaufman Speech Praxis Test

The Kaufman Speech Praxis Test (KSPT) was administered to all participants to evaluate vocal imitation of consonants, vowels, and syllable shapes of increasing complexity (Kaufman, 1995a). While developed to aid in the diagnosis and treatment of childhood apraxia of speech, the KSPT is a standardized, norm–referenced assessment that offers several advantages over other tests of articulation when evaluating vocal imitation. For one, only a verbal model is presented, and the child's vocal response is recorded. There are no additional verbal stimuli or pictures, thus making for a simple, pure measure of echoic behavior. Test items may be repeated as needed. The KSPT is divided into four parts (Kaufman, 1995b). Part 1 evaluates oral movements, such as open mouth, lateralized tongue movements, and rounded lips. Part 2 assesses imitation of early developing phonemes (e.g., vowels, nasals, bilabial and alveolar plosives) in simple syllables. Part 3 evaluates later developing phonemes (e.g., velar plosives, fricatives, affricates, glides) and more complex syllables such as bisyllabics and polysyllabics. Part 4 examines length and complexity of spontaneous speech. Each part of the KSPT is standardized and separately norm–referenced for both typical (i.e., not identified as speech impaired) and disordered (i.e., identified as speech impaired) populations 24–72 months of age. Consequently, the specific part(s), as appropriate to the child's current level of functioning, may be administered and scored without unnecessary probing of developmentally inappropriate vocalizations (Kaufman, 1995a).

Only parts 2 and 3 of the KSPT were considered for the participants of the present study because these sections measure echoic behavior. Part 2 was administered to all participants in the current study. Part 3 was only administered, per KSPT guidelines, if the child had few errors in part 2 and the child had a significant length of utterance defined as 4+ word combinations in spontaneous speech (Kaufman, 1995b).

Parts 2 and 3 were administered to Sarah. She correctly imitated all pure vowels and simple consonants in isolation. While most diphthongs were correctly produced, distortions were observed on /ou/ and /eɪ/. Consonant errors at the syllable level included /n/ as a substitute for /m/ in initial and medial positions (e.g., “mama” pronounced /nɑnɑ/) and /p/ as a substitute for /m/ and in final positions (e.g., “home” pronounced /hop/). Consonant /p/ was also observed as a substitute for /n, b, d/ in the final position of CVC words, although inconsistently. Because Sarah reportedly produced 3–4 utterances on occasion, part 3 of the KSPT was also administered. She correctly imitated /k/ and /g/ in isolation and maintained production in initial and final positions of CVC syllables. Sarah also imitated /f, w, j, l, r, s, z, ʃ/ in isolation, although errors were observed on these sounds in initial and medial positions of CVC syllables. Numerous errors were observed across s–blends, r–blends, and l–blends. Simple cluster reductions (e.g., “skate” pronounced /keɪt/) and cluster substitutions (e.g., “stop” produced as /dɑp/) were noted, while most consonant blend productions contained multiple errors, including final consonant deletion, voicing, assimilation, and gliding. All CVC syllables requiring either front–to–back or back–to–front synthesis were correctly imitated (e.g., “dig,” “cat”). Additionally, Sarah was able to imitate some complex bisyllabics and polysyllabics such as “wagon” and “banana.” Due to Sarah’s chronological age, standard scores and percentile rankings could not be calculated on the KSPT. However, the number of errors and her overall raw score on part 2 and part 3 earned an age equivalency of < 2.0 years.

Jayden received part 2 of the KSPT. He correctly imitated simple phonemes /ɛ/ and /m/. Jayden attempted to imitate all other modeled sounds and syllables although productions were incorrect. Most vocalizations were the open vowel sound /ɑ/, open–mid vowels /ɛ, ʌ/, and the consonant /m/. When presented /h/ in isolation and words such as “hot,” “hop,” and “happy,” Jayden produced the /h/ followed by a vowel sound (e.g., /hɛ/, /hɑ/). A standard score and percentile ranking could not be calculated according to Jayden’s chronological age. The number of errors and raw score for part 2 generated an age equivalency of < 2.0 years.

Part 2 of the KSPT was administered to Benjamin. All sounds, syllables, and words were presented, although attempts to imitate were rare. The only vocalizations observed during the assessment were correct imitation of /n/ in isolation and /ʌdə/ following presentation of “papa.” Standard scores and percentile rankings based on normative data from the typical population could not be obtained. However, the number of errors and raw score generated an age equivalence of < 2.0 years. According to normative data from the disordered population, Benjamin earned a standard score of 46, placing him in the 3rd percentile.

Theo received part 2 of the KSPT. Errors were observed on pure vowels /u, ɔ/ and diphthong /ɔɪ/. Production of consonant /p/ was either voiced (i.e., /b/) or followed by an additional vowel (e.g., /pə/), and these errors were observed throughout part 2. Imitation of other simple consonants in isolation were occasionally followed by a schwa vowel sound, but Theo demonstrated correct imitation on repeated attempts. All CVCV reduplicated syllables were correctly imitated, with the exception of /pɑpɑ/ which was marked by voicing of the consonants (i.e., /bɑbɑ/). Voicing of consonants /p/ and /t/ were observed during imitation of CV syllables. Theo had difficulty imitating all VCV syllables with most errors on the final vowel and voicing /p/. Final consonant deletion was observed on attempted imitations of CVC syllables with both unchanging consonant sounds. After repeated attempts Theo imitated “home” with /m/ in the final position, but all other productions of CVC syllables with changing consonant sounds were characterized by final consonant deletion. Some productions of simple CVCV bisyllabics with changing vowel and consonant sounds were marked by assimilation processes such that the consonants remained unchanged (e.g., “bunny,” pronounced as /bʌbi/). Based on normative data from the typical population, Theo earned a standard score of 15, placing him in the 2nd percentile. The number of errors and raw score generated an age equivalence of < 2.0 years. According to normative data from the disordered population, Theo earned a standard score of 71, placing him in the 8th percentile. Part 3 was not administered because multiple errors were observed on part 2 and Theo’s therapists and parents reported utterance length in spontaneous communication was typically 1–2 words.

A summary of participant characteristics can be found in Table 1.

Table 1

Summary of Participant Characteristics

Participant	Age (years)	Communication Modality	Mean BLAF Score[a]	BLAF Motor Imitation Score[a]	BLAF Echoic Score[a]	KSPT Part 2 Standard Score[b] (D)	KSPT Part 2 Standard Score[b] (TD)	KSPT Part 2 Age Equivalency (TD)
Sarah	10	SGD	4.6	5	4	N/A	N/A	< 2.0 years
Jayden	11	SGD	2.4	2	1	N/A	N/A	< 2.0 years
Benjamin	5	Signs	1.8	3	1	46	N/A	< 2.0 years
Theo	3	Vocal	2.2	4	2	71	15	< 2.0 years

Note. SGD = Speech generating device; BLAF = *The Behavioral Language Assessment Form;* KSPT = *Kaufman Speech Praxis Test;* (D) = Normative data from disordered population; (TD) = Normative data from typically developing population.

[a] Informant assessment scored on a scale of 1–5

[b] Norm referenced assessment based on a mean score of 100 with a standard deviation of +/– 15

Stimulus Preference Assessment

To identify preferred items and/or activities, a modified version of the *Reinforcer Assessment for Individuals with Severe Disability* was used to interview parents (Fisher et al., 1996). The interview questions divide potential preferences into a range of categories including visual, auditory, olfactory, edible, and tactile stimuli. Additional areas include preferences associated with movement, temperature, attention, and toys. Thus, the questions are intended to yield a more comprehensive and individualized list of preferred stimuli than less formal, open–ended interviews. The results of the structured parental interview were briefly reviewed with each participant's therapy team to confirm reported preferences and inquire further about preferred stimuli specific to school and clinic environments, of which parents may have been unaware.

Preferred stimuli identified through the interview were then presented in a brief multiple stimulus without replacement (MSWO) array. The brief MSWO was conducted in accordance with the procedures described by Weaver et al. (2017). Items were presented in a straight line, spaced approximately 2–4 inches apart. When possible, one small sample of the edible item was presented (e.g., one cracker, one yogurt melt). Otherwise, the entire package was presented (e.g., juice box, yogurt cup). Participants were led to the table and offered a simple question/instruction such as, "Pick one" or "What would you like?" or similar phrase. After selection, the participant was gently led away from the table and allowed to consume the edible or play with the tangible for approximately 20 s. Edibles presented in their packaging were removed after a couple of sips/bites. The remaining items were rotated and equally spaced apart. The procedure was repeated until all items were selected or until the participant refused the remaining items. A total of three MSWO assessments were completed. Data were collected on the order of items selected at each assessment. The percentage of selections was calculated by adding the total number of selections divided by the total number of presentations and multiplying by 100. Parents and therapists were asked to provide the five highest ranked items for use during treatment sessions.

According to Sarah's mother, Sarah enjoyed visual stimuli such as looking at herself in the mirror, viewing shows on YouTube, watching the movie *Trolls*, looking at the movement of spring toys, and playing video games. Music was noted to be a highly preferred auditory stimulus including instruments and songs from popular children's movies. Sarah was described as a picky eater, but enjoyed a few snack items such as yogurt melts, cheese crackers, drinkable yogurt, and marshmallows found in a common breakfast cereal. Bouncing was also reported as a highly preferred activity especially with items, such as pogo sticks, physiotherapy balls, and the trampoline. Other movement activities included swinging, climbing, riding a bike, and occasionally being chased. Tactile sensations, such as fidget toys, modeling dough, sand, and dried rice, were reported to be enjoyable. Toy preferences were more limited but included dolls representing characters from favorite television shows and movies, as well as those that make noise. Sarah reportedly enjoyed most items and activities alone, although her mother explained that she valued verbal praise. When asked what Sarah would likely do if given free access to her favorite items and activities, her mother anticipated that she would play on her tablet or video games on her Wii.

The results of the structured parental interview were briefly reviewed with Sarah's therapy team to confirm reported preferences and inquire further about preferred stimuli specific to school and clinic environments of which Sarah's mother may have been unaware. The team suggested fidget toys/items such as therapy putty, a sensory brush, a rubber noodle, and a book of coloring sheets depicting characters from one of Sarah's favorite television shows that were only available at the clinic. To avoid disruptions during treatment, Sarah's therapy team recommended that electronic devices (e.g., tablet, phone) not be used to reinforce responses even though they were highly preferred. Removing these items would potentially evoke problem behaviors. Edibles provided an advantage in the use of discrete trial instruction. As such, the stimuli presented in the MSWO assessment included yogurt melts, two brands of cheese crackers, juice, drinkable yogurt, therapy putty, sensory brush, coloring book, and rubber noodle. The yogurt melts, both types of cheese crackers, juice, and drinkable yogurt ranked highest. At the beginning of each session 3–5 of the highest ranked items were presented to Sarah, and the item selected was used to reinforce correct responses during treatment. She most frequently selected cheese crackers and occasionally juice or yogurt melts.

Jayden had a history of pica, and while his treatment had reduced these behaviors, his mother urged close supervision with the use of any inedible items. Hence, stimuli associated with taste were highly preferred. Some of Jayden's favorite foods were those with more intense flavors and textures such as basil, strong peppermints, lettuce, kale, salt, and carbonated drinks. Preferred movements included swinging, rocking back and forth, and swimming. Jayden's mother explained that Jayden often throws or drops items on hard surfaces, although she attributed this preference more to the auditory sensation than the movement. Dropping items produced a louder sound on contact, which Jayden seemed to prefer. Other auditory stimuli such as music, echoes, sound reverberations, and louder sounds in general were also preferred. Visual stimuli seemed less preferred, as he occasionally enjoyed playing a game on his tablet. Jayden's mother explained that Jayden liked cold temperatures and was usually very excited when he walked outside on a cold morning. Tactile sensations such as fuzzy blankets, dried rice, and heavy mats were preferred. Toy preferences included simple cause/effect and musical toys. Jayden also enjoyed fidgeting (i.e., dangling, shaking, flicking) with toys and objects such as springs, ribbons, and hard plastic items. Although he reportedly enjoyed tickles and pats on the back from certain people, Jayden mostly preferred to play alone. When asked, hypothetically, what Jayden would do if he were given unrestricted access to his favorite items and activities, his mother expected that he would either find something to eat, or walk around fidgeting with a keychain, small plastic shovel, spring, or other similar item.

Jayden's therapy team confirmed his preferences for edibles. Based on the reported preferences and availability of edibles, strong peppermints, lollipops, fruit flavored candy, basil, macaroni noodles, and lemon–lime soda were presented in the MSWO. The items were ranked as follows: fruit flavored candy, soda, basil, macaroni noodles, and lollipops. At the start of each session, 3–5 of these items were presented to Jayden and the selected item was used to reinforce correct responses during treatment. However, Jayden often refused each of the presented items either at the start of the session or after selection, when the item was delivered in treatment. His therapy team explained that a variety of foods were noncontingently available to Jayden throughout the day to reduce pica related behaviors. These procedures appeared to have had an abolishing effect on the preferred edibles. To identify preferred stimuli to be delivered contingently throughout the sessions, Jayden was simply asked what he would like to work for and verbally prompted to use his device to indicate his preference. Alternatively, he was presented with preferred tangibles (i.e., fidget toys) or taken to an area of the clinic and instructed to pick item(s) that he would like to earn. His requests and/or selections were confirmed by response to a yes/no question (e.g., "Is this what you want to work for?" followed by a head shake/nod). Most frequent selections were arbitrary items that may best be described as fidget toys (e.g., block, shoestring, plastic tube, stretchy band).

Benjamin’s preferences were mostly limited to taste, movement, touch, attention, and a few toys. He liked chicken nuggets and fries, and his mother also reported a preference for crunchy foods such as chips, popcorn, and cheese crackers. Benjamin enjoyed movement activities such as swimming, swinging, flapping his hands, bouncing on the trampoline or bed, and riding in the golf cart. Tactile preferences were associated with those of movement, as Benjamin liked to splash in the water and feel wind blowing on his face. He reportedly enjoyed verbal praise and hugs and showed an interest in toys that light up and make sound. When asked, hypothetically, what Benjamin would do if he were granted free access to his favorite items and activities, his mother predicted that he would sit and eat a box of cheese crackers.

Benjamin’s therapy team confirmed his preference for salty snacks and also suggested a rubber noodle that he was observed to play with frequently. Thus, items presented in the MSWO included popcorn, wheat crackers, two types of cheese crackers, potato chips, pureed fruit, and a rubber noodle. The rubber noodle, both types of cheese crackers, potato chips, and wheat crackers ranked highest. At the beginning of each session 3–5 of the highest ranked items were presented and the item selected was used to reinforce correct responses during treatment. The therapy team later reported that Benjamin developed a preference for chocolate candies. These were also included in the presession choice when available. Benjamin most frequently selected the rubber noodle, cheese crackers, and chocolate candies.

According to Theo’s father, Theo had a high preference for visual stimuli, particularly items that spin, including ceiling fans, pinwheels, windmills, and propellers. By extension, images and toy replicas of these objects were also preferred, such as pictures and videos of ceiling fans, toy helicopters, toy airplanes, and fidget spinners. The children's movie, Encanto, was reported to be his favorite. He enjoyed singing, although loud sounds such as those produced by a lawn mower or blender were aversive. He reportedly ate a variety of foods. Specific taste preferences included gummy bears, cookies, chocolate candies, cheese puffs, and cheese crackers. Movement activities such as running, and climbing were also preferred. While Theo did not like rain or water from the shower falling on his head, he did enjoy other tactile sensations, such as splashing in the bath, feeling the air from a fan, and being tickled. He reportedly liked verbal praise from others and would give hugs and high fives. When Theo's father was asked what he thought Theo would do if given unrestricted access to his favorite items and activities, he stated that Theo would most likely move about the house turning the lights on and off, turning the ceiling fans on and off, and opening and closing doors.

Theo's parents provided preferred edibles, and his therapy team helped identify preferred tangibles available at the clinic. Items presented in the MSWO included: gummy bears, cheese crackers, cheese puffs, lollipop, tablet (with *Encanto*), wristband (with propeller), toy windmill, pinwheel, toy airplane, and puzzle. The highest ranked items were the lollipop, toy airplane, cheese crackers, tablet, windmill, and gummy bears. At the beginning of each session, 3–5 of the highest ranked items were presented, and the item selected was used to reinforce correct responses during treatment. Theo's parents sent chocolate candies later during treatment, and these were included in the pre–session choice, when available. Theo most frequently selected lollipops and chocolate candies. Over time, he spontaneously requested his desired item, before items were presented in the presession choice. These vocal requests were accepted, and the item was offered. Theo's speech intelligibility was poor, so his vocal requests were confirmed by response to a yes/no question (e.g., “Is this what you want to work for?” followed by a head shake/nod and vocal approximation).

Experimental Design

A multi–element design was used to compare two independent variables—conventional vocal imitation training and vocal imitation training with physical prompts—on an echoic response. One vocal target was assigned to each treatment condition. The dependent variable was correct imitation on two pre-session echoic probes of each vocal target. After reaching the mastery criteria for a vocal target (i.e., correct imitation on both presession echoic probes for five consecutive sessions), the treatment was discontinued. Maintenance of the echoic was evaluated 7–15 days later with two echoic probes.

Target Vocalizations

Several variables were regarded in the selection of target vocalizations. First, the target words, sounds, and syllables could not be addressed in other therapies. Second, developmental norms for speech acquisition were considered (see Dyer, 1987). Earlier developing sounds and syllable shapes were targeted before later developing sounds as long as they were not already within the child’s repertoire. Chronological age was also a factor, such that only those sounds expected to be mastered were targeted. For example, an 8–year–old child would typically have acquired /m/, whereas a 3–year–old would not be expected to master production of /r/ (Roth & Worthington, 2018). Third, parent input was considered. When asked, parents expressed preferences for target vocalizations, and these requests were honored if appropriate to the child’s developmental level. Fourth, because echoics are often used in transfer procedures, the potential for the vocal to function as another verbal operant was considered. Finally, two target vocalizations were selected for each participant (one per condition), and it was necessary that these targets matched in terms of developmental acquisition and functional use. Yet, it was also important that the two targets differed in their acoustic properties and production features (e.g., place, manner, voicing, jaw height, tongue position) to facilitate discrimination. For example, /b/ and /p/ are well matched in terms of age of acquisition and ease of production, but differ only in voicing, thus making discrimination difficult.

Considering Sarah's chronological age and the developmental norms for speech acquisition, full mastery of all phonemes and "adult–like" speech would be expected. Her mother mentioned Sarah's difficulty with voiced and voiceless "th" sounds and possibly targeting those in treatment. However, based on Sarah's performance on the KSPT, she produced many sounds in isolation that were not maintained across syllables. For example, Sarah produced /m/ (a very early developing sound) in isolation, but at the syllable level, she substituted /n/ for /m/. Therefore, addressing these errors at the syllable level would have a greater impact on her overall speech intelligibility. As such, CVCV syllables with changing consonant and vowel sounds were selected but designed to include the consonants observed in error. The words "movie" (/mʌvi/) and "funny" (/fʌni/) were selected and randomly assigned to the treatment conditions. Neither syllables, nor their individual phonemes, were being targeted in other therapies. Both targets contained nasal sounds, labiodental fricatives, and identical vowels and syllable shapes, but were also easily discriminable. Based on Sarah's language it was anticipated that, upon mastery, these targets could be transferred to mands and tacts.

Given Jayden's chronological age, full mastery of all phonemes and "adult–like" speech would be expected. His mother suggested several sounds and syllables that could later be transferred to mands or intraverbals including /no/ or /nɑ/ as an approximation or word for "no", /mɑ/ or /mɑm/ for his mother, /o/ as an approximation for "open", and /b/ as an approximation for “bathroom.” Responses on the KSPT were mostly open to mid–range vowel sounds such as /ɑ, ʌ, ɛ, ə/. When sequenced with consonants, the resulting syllable shapes would likely become functional word approximations. Therefore, consonants /b/ and /n/ were selected and randomly assigned to the treatment conditions. Neither of these sounds were being targeted in other therapies. Both sounds are voiced and are typically acquired around 3 years of age (Roth & Worthington, 2018), but they differ in place and manner of articulation.

Benjamin’s mother suggested an approximation for “more” and the sound /i/, as an approximation for “eat”. During the KSPT assessment, Benjamin rarely vocalized in response to the modeled sounds. Vowel–like sounds were most often observed in spontaneous productions. Thus, to teach vocal imitation, vowels /ɑ/ and /i/ were selected as targets and randomly assigned to treatment conditions. Neither of these sounds were being targeted in other therapies. Because Benjamin was potty training, it was expected that /ɑ/ could later function as a mand for the restroom (i.e., /ɑ/ as in /pɑti/ or “potty”), while /i/ could function as a mand for “eat.” Although both targets were vowels, they differed in jaw height, tongue position, and acoustic features.

Theo’s father expressed concerns regarding Theo’s final consonant deletion and production of /s/. Common errors on the KSPT included omission of final consonants in CVC syllables with changing and unchanging consonant sounds. When presented with consonant sounds in isolation, the occasional addition of a schwa–like vowel sound was also noted (e.g., pronounced /d/ as /də/). Final consonant deletion is typically suppressed by 36 months (Bauman–Waengler, 2004). To improve speech intelligibility and address parental concerns, two VC targets were selected, /it/ and /ʌp/. Neither syllables, nor their individual phonemes, were being targeted in other therapies. Given Theo’s preference for ceiling fans and perseveration on this topic during interactions with others, it was expected that “up” would have many opportunities for functional use and “eat” could function as a mand, tact, or intraverbal. Each of these individual phonemes were correctly produced in other contexts. Both targets have VC syllable shapes and voiceless, stop consonants that typically emerge between 3–4 years of age (Roth & Worthington, 2018). However, the vowel sounds are dissimilar, and the consonants differ in placement, allowing for easier discrimination.

Procedures

Echoic Probes

Due to the simultaneous prompts used in treatment, pre-session echoic probes were necessary dependent measures of vocal imitation (see Leaf et al., 2022). Prior to each session, two probes of each target vocalization were conducted. In the probe trial, the therapist modeled the vocalization (e.g., “Say, movie”, “movie”) and recorded the participant’s response. Participants had approximately 3–5 s to respond before the next probe was presented. No programmed consequences followed responses. That is, no praise or access to preferred stimuli was provided following correct or incorrect responses. Echoic probes were scored according to the criteria developed by Civindi–Motta et al. (2017), which defined a correct response as, “any vocalization that occurred within 5 s of the sample vocalization and matched the sample vocalization in number of phonemes as well as auditory similarity, excluding vocal stereotypy” (p. 46). Slight errors in placement or other atypical articulatory patterns were disregarded if the phoneme was acoustically the same as the model. For example, Sarah was observed to occasionally produce "ee" with jaw sliding and unilateral lip retraction. Benjamin exhibited slightly reduced or overextended jaw height on production of “ah.” Nonetheless these responses were counted as correct.

Two probes of one vocal target were presented consecutively before probing the other vocal target, but the order of the two targets was semi–random. For example, the two probes for "movie" were conducted before the two probes for "funny" were presented. On subsequent sessions the two probes for "funny" may have preceded the two probes for "movie." Typically, the order alternated every 2–3 sessions.

Treatment

At the beginning of each session 3–5 of the highest ranked items were presented to participants, and one of the items selected was used to reinforce correct responses during treatment. Chocolate candies were added to Benjamin's and Theo's pre-session choice when available. Jayden was also permitted to indicate his preference via his speech generating device or select a tangible from available toys. Each treatment condition, vocal imitation training and physical prompts, consisted of 20 mass trials of the target vocalization assigned to that condition. Initially, in discrete trial teaching, new targets are often addressed through mass trial presentations (see Eikeseth et al., 2014; Grow et al., 2011). Although few studies have compared mass trial and interspersed or random trial presentations, particularly in echoic training, Henrickson et al. (2015) found that the rate of acquisition was higher for mass–trial instruction than interspersed–trial teaching. As such, and in accordance with mass–trial teaching, a total of 20 mass trials were conducted to allow for ample repetition and rehearsal while also minimizing inattention and other unwanted behaviors (see Henrickson et al., 2015). The order of the two treatment conditions at each session were semi–randomly rotated, with the order usually changing every 2–3 sessions. Praise was typically provided every 1–2 min, for behaviors such as remaining seated, looking at the therapist, attempting to imitate, and cooperating (e.g., "You're working so hard!," "Yes, you're sitting perfectly. Let's do some more."). Every 3–4 min the therapist offered a break away from the table where the participant could stand, move about the room, and interact with other items. Breaks usually lasted 1–2 min although they were longer if needed (e.g., restroom, transition). Overall, sessions were conducted once daily, 4–5 times per week, and lasted approximately 20 min in duration.

Vocal Imitation Training. During the vocal imitation training condition, the therapist encouraged the participant to attend (e.g., positioned the participant's chair to face the therapist, asked "Okay, ready?"), modeled the vocal target, and instructed the participant to imitate (e.g., "Say, /ɑ/," "Look at me, /ɑ/"). Correct responses were followed by verbal praise (e.g., "Yes! That's it!" or "Good work!") and access to the preferred item or edible selected by the participant at the start of the session. After the edible was consumed, or the participant engaged with the tangible for 5–10 s, the next trial was presented. To mimic procedures used in clinical applications, the therapist often provided gestural prompts (i.e., pointing at her face/mouth) and slightly exaggerated model which modified features of the stimulus (e.g., duration, magnitude) to increase its saliency. For example, when modeling /i/ for Benjamin, lip retraction was more pronounced and the vowel sound was elongated (i.e., "eeee"). If the participant did not respond on the first presentation, or looked away after the model, the trial was presented again. A shaping procedure was used with Sarah, Benjamin, and Theo in which the full vocal target was reduced to single sounds, partial syllables or segments, or simpler articulatory movements (i.e., fewer articulatory components) and successive approximations toward the full sound/syllable were reinforced. With Sarah and Theo, simultaneous vocal prompts were provided, such that the therapist and participant vocalized together. This was attempted with Jayden and Benjamin, although they did not produce the modeled sound.

Physical Prompts. The procedures for the physical prompts condition were like those in the vocal imitation training except that physical prompts were applied to the face and mouth. The therapist encouraged the participant to attend (e.g., positioned the participant's chair to face the therapist, asked "Okay, ready?"), modeled the vocal target (e.g., "Say, /ɑ/," "Look at me, /ɑ/"), and immediately provided the physical prompt. While administering the physical prompt, the therapist simultaneously produced the vocal target. Correct responses were followed by verbal praise (e.g., "Yes! That's it!" or "Good work!") and access to the preferred item or edible selected by the participant at the start of the session. After the edible was consumed, or the participant engaged with the tangible for 5–10 s, the next trial was presented. If the participant was ill positioned for the physical prompt, the therapist repeated the trial. For example, during initial sessions Sarah lifted her chin toward the therapist, anticipating touch to her face and mouth. Benjamin occasionally turned his head to the side or downward. In these cases, the therapist repositioned their head, modeled the sound, and applied the physical prompts while vocalizing the target sound/syllable.

The specific physical prompts used were individualized to the participant's current vocal responses and selected echoic targets. Sarah's target vocalization for the physical prompts condition was /muvi/ (i.e., "movie"). The therapist first supported Sarah's head by gently placing her left forearm between Sarah's right shoulder and head while placing her left palm at the back of Sarah's head to maintain an upright, neutral head position. Using a parameter prompt for /m/ from the PROMPT approach, the therapist placed her right index finger along Sarah's jaw line, her bent right middle finger under the chin, and the back of her right thumb pressed lightly on the lips (see The PROMPT Institute, 2009). While a surface prompt may have been more suitable to teach the co–articulatory movements in the target word, the parameter prompt provided input to both the jaw and lips and was selected because Sarah typically emitted /n/ in place of /m/ at the syllable and word level. A surface prompt was used for production of /u/. With this prompt, the therapist placed her right thumb and index finger at the labial corners, and gently applied light pressure and horizontal, inward movement toward the therapist to promote lip protrusion or puckering. To prompt /v/ using a surface prompt, the then placed her right thumb on or just below the lower labial surface and provided moderate inward pressure. A surface prompt for /i/ included placement of the right index and thumb at the labial corners, with light pressure and outward movement away from the therapist to promote lip retraction (e.g., as if forming a smile).

Jayden's target vocalization for the physical prompts condition was /n/. The therapist used supports similar to those described with Sarah to maintain an upright, neutral head position. A modified surface prompt from PROMPT was provided (see The PROMPT Institute, 2009). Placing her right middle finger on the mylohyoid muscle, just posterior to the mandible, the therapist provided a firm upward movement, while providing light touch to nasal ala (i.e., wing of the nose just outside the nostril) with her right thumb. The firm pressure on the mylohyoid was to promote tongue elevation by providing a kinesthetic cue to the muscle, and touch to the nasal ala was intended to cue the nasal airflow. These two movements, tongue elevation and nasal airflow, are necessary for production of /n/. Because Jayden often produced /m/ when /n/ was modeled, the therapist often placed her thumb anteriorly on his chin to slightly lower the position of the jaw before providing touch to the mylohyoid.

Benjamin's target vocalization for the physical prompts condition was /ɑ/. The same supports as previously described were used to maintain an upright, neutral head position. Components of both the complex and parameter prompt from the PROMPT method were applied (see The PROMPT Institute, 2009). The therapist placed her right index finger along Benjamin's jaw line, her bent right middle finger under his chin, and her right thumb on the anterior surface of his chin to guide the mandible to an open position, gently pressing in a downward motion. Initially, Benjamin resisted the movement facilitated by physical prompts. Clenching, jaw sliding, and jaw jutting occurred frequently and so a slight, but unsuccessful modification to the prompt was provided intermittently over the first week of treatment. The therapist moved her right thumb from the anterior surface of the chin to the jaw, to provide broader support along the jawline. Because the same resistance to this prompt was observed, the original prompt, with the right thumb placed on the chin, was used. Later, a final component was added to prompt the feature of voicing. The therapist used the same physical prompt and additionally placed her right second finger gently on the side of the larynx to cue the vibration necessary for voicing.

Finally, Theo's target vocalization for the physical prompts condition was /it/. The same supports as previously described were used to maintain an upright, neutral head position. Modified surface prompts from PROMPT were used to teach imitation of both phonemes in the syllable target, namely /i/ and /t/. For /i/, the therapist placed her thumb and first finger at or just above the labial corners, with a light outward movement away from the therapist to promote lip retraction (e.g., as if forming a smile; see The PROMPT Institute, 2009). Placing her right middle finger on the mylohyoid muscle, just posterior to the mandible, the therapist provided a firm upward movement to promote tongue elevation necessary for production of /t/ by providing a kinesthetic cue to the muscle responsible for lingual elevation.

Social Validity

At the conclusion of the intervention, a questionnaire was emailed to behavior technicians and parents of participants to assess social validity. The behavior technicians had directly observed the treatments and responded based on those observations. Since parents were unable to observe the treatments, video clips were provided, which depicted: 1) baseline echoic responses for both vocal targets, 2) clips from several vocal imitation training sessions and physical prompts sessions, and 3) post–treatment echoic responses for both vocal targets. Responses to the questionnaire that depicted perceptions and opinions of the treatment conditions were reported according to a 5–point Likert scale of agreement (*strongly agree* to *strongly disagree*). Additionally, two opportunities for open–ended comments on each treatment condition were included.

Results

Data on the echoic probes for each treatment condition are presented in Figures 5, 6, 7, and 8. All participants acquired new echoic responses as a result of the interventions. Sarah and Theo mastered target vocalizations under both treatment conditions. Interestingly, Jayden and Benjamin learned only the echoics taught using physical prompts.

Sarah

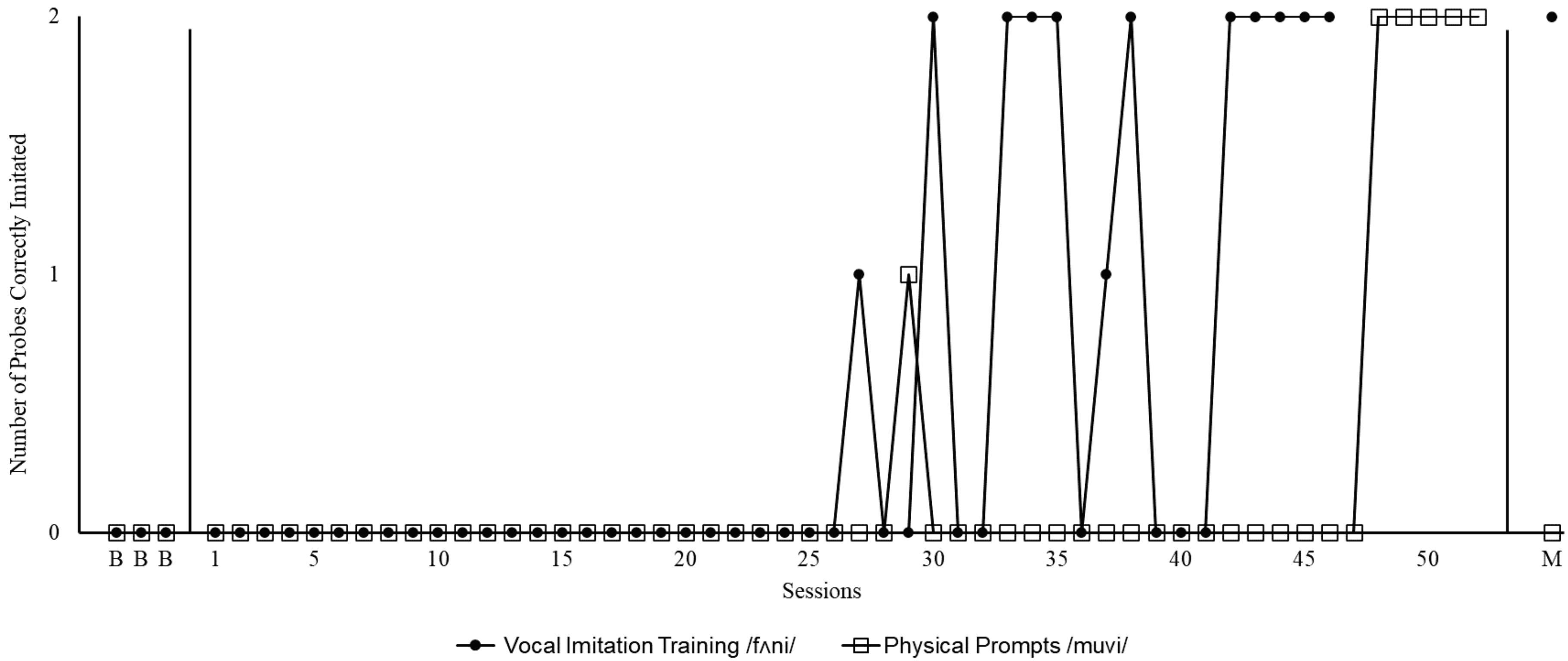
Sarah's target vocalization for the vocal imitation training condition was /fʌni/ (i.e., "funny"). Errors were typically /fbʌni/ or /bʌni/. To teach sequencing, a shaping procedure was introduced on the second treatment session. Sarah correctly imitated /f/ in isolation, so the session proceeded with the target /fʌ/. However, errors persisted; she had difficulty with speech motor sound transitions across consecutive phonemes such as /f/ to /ʌ/ and often inserted a stop which resulted in productions like /fbʌ/. As a result, /f/ in isolation was targeted over the next two sessions though Sarah easily imitated the sound. To teach the speech motor sound transition from /f/ to /ʌ/, the CV syllable was modeled, and then the therapist used a simultaneous vocal prompt to evoke the correct response. After modeling the syllable /fʌ/, the first phoneme of the syllable, /f/, was slowly presented again, and Sarah produced the sound in unison with the therapist. The second phoneme /ʌ/ was presented in the same way and Sarah produced the sound in unison with the therapist. Over time the therapist produced this simultaneous vocal prompt with a gradually reduced volume until the prompt only consisted of the inaudible articulatory movements. Using the same simultaneous vocal prompting procedures, /fʌn/ was targeted. Sarah began adding a schwa vowel sound to the syllable (i.e., /fʌnə/). In an effort to interrupt this response, the therapist provided praise and reinforcement immediately after production of /n/, but the error persisted. Treatment then focused first on imitation of /n/ in isolation, followed by the syllables /ʌn/ and /ʌni/. Finally, the full vocal target /fʌni/ was addressed with the same simultaneous prompting procedures. Sarah correctly imitated /fʌni/ during treatment with the continuous schedule of reinforcement, but not consistently in the presession echoic probes. Therefore, the schedule of reinforcement was reduced to a variable ratio 2, in an effort to strengthen and stabilize responding.

The target vocalization for the physical prompts condition was /muvi/ (i.e., “movie”). Errors were typically /nuvni/ or /fnuni/. As in the vocal imitation training condition, a shaping procedure was introduced on the second treatment session. Sarah correctly imitated /m/ in isolation, so the session proceeded with the target /mu/. Physical prompts were immediately administered for each phoneme while the therapist simultaneously vocalized. Again, Sarah exhibited difficulty with speech motor sound transitions across consecutive phonemes. When transitioning from /m/ to /u/ she would typically insert an additional sound between the phonemes, producing /mnu/. Treatment then focused on correct imitation of /m/ in isolation. Again, like in the vocal imitation training condition, Sarah easily imitated the initial consonant, /m/, so treatment proceeded with imitation of /mu/. Immediately after modeling the syllable, the therapist provided the same physical prompt for /m/ and a surface prompt for /u/ with simultaneous vocal production. Next, imitation of /muv/ was taught following the same procedures. The full target vocalization /muvi/ was introduced, but Sarah began inserting /n/ before the final vowel such that her production was /muvni/. Correct production of the syllable /vi/ was addressed with physical prompts and simultaneous vocal production so that Sarah learned to transition from /v/ to /i/. The full vocal target was reintroduced with physical prompts and simultaneous vocalizations for each phoneme. The physical prompts were faded according to a backward chaining approach. That is, the therapist physically prompted every phoneme in the syllable sequence but omitted the prompt for /i/. Next, the physical prompts were provided for all phonemes except /v/ and /i/. The physical prompt for /m/ faded by gradually increasing the proximity between the therapist's hand and Sarah's face until only "air prompts" were provided (Hayden, 2004; L. Slim, personal communication, 2022). When Sarah incorrectly imitated /muvi/, the therapist provided the physical prompt for /m/, to evoke the correct production. Correct imitation of /muvi/ was observed during treatment but not in the presession echoic probes. As a result, the reinforcement schedule was reduced to a variable ratio 2, to strengthen and stabilize responding.

Data for Sarah's imitation of echoic probes is displayed in Figure 5. She acquired imitation of /fʌni/, taught in the vocal imitation training condition, before mastering /muvi/ in the physical prompts condition. After 26 treatment sessions, Sarah correctly imitated /fʌni/ on one echoic probe. Performance fluctuated thereafter and, on session 46, Sarah demonstrated mastery of the echoic target /fʌni/ with correct imitation on two probes for five consecutive sessions. Total training time to reach the mastery criterion for /fʌni/ was 900 trials over 45 treatment sessions. Treatment was discontinued, and 15 days later correct imitation was maintained across two echoic probes. Correct imitation of /muvi/ first occurred after 28 treatment sessions. Errors continued until session 48 when Sarah correctly imitated both probes. Performance stabilized thereafter, and Sarah mastered imitation of /muvi/on treatment session 52. Total training time to reach the mastery criterion for /muvi/ was 1,020 trials over 51 treatment sessions. Maintenance probes were conducted 7 days later but Sarah did not correctly imitate the target on either probe. Instead, she emitted her original errored production /fnuni/ on both probes.

Figure 5

Echoic Probes for Sarah



Note. The number of correct imitations across two presession echoic probes. B = Baseline; M = Maintenance

Jayden

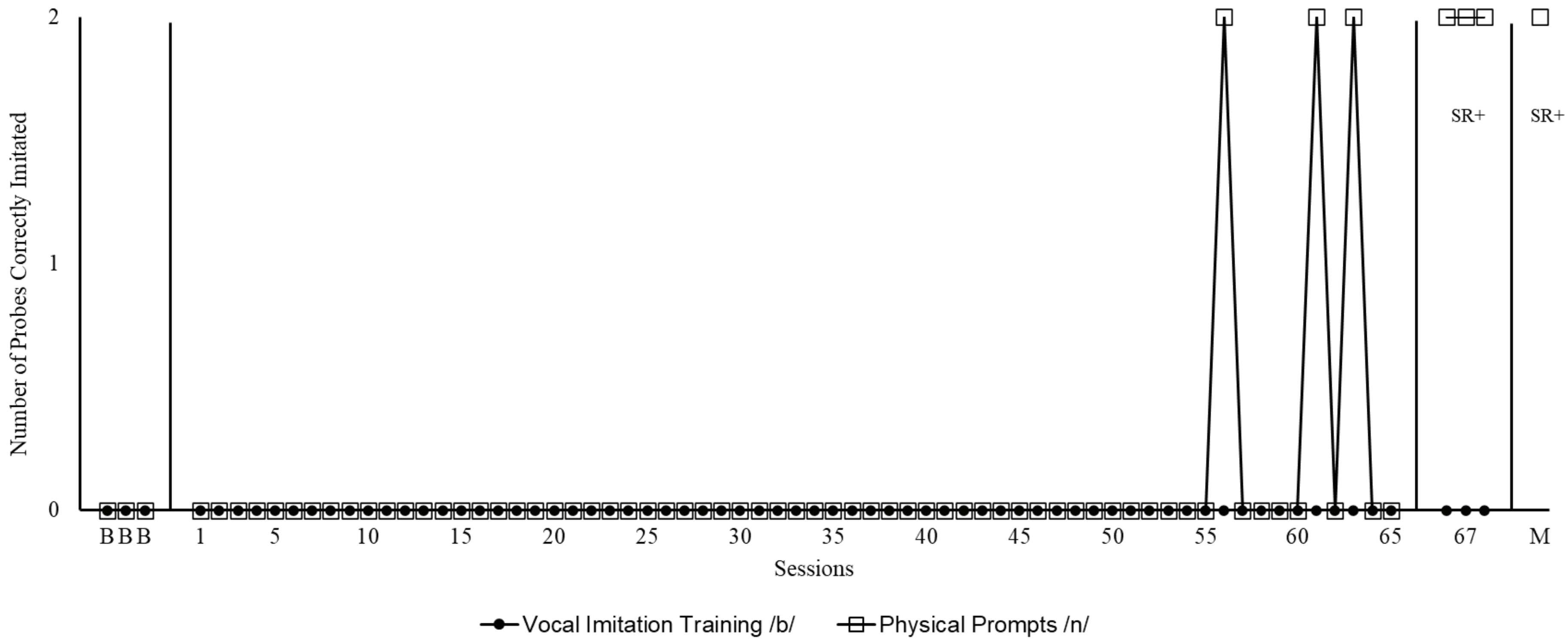
Jayden's target vocalization for the vocal imitation training condition was /b/. Errors were always production of /m/ or an open to mid vowel sound such as /ʌ/, /ɛ/, /ə/, or /ɑ/ (as is the first vowel sound in "up", "echo", "about", and "olive", respectively). Occasionally, the vowels were preceded by /j/ (as is the first consonant sound in "yellow") to produce /jʌ/ or /jɛ/. Additional verbal instructions were intermittently provided before the model such as "Watch my lips," or "Do it like me," while pointing toward the lips/face. Jayden never approximated or imitated the target sound.

The target vocalization for the physical prompts condition was /n/. Again, errors were always production of /m/ or an open to mid vowel sound such as /ʌ/, /ɛ/, /ə/, or /ɑ/. The physical prompt for /n/ was sequentially faded by removing touch to the nasal ala, providing less pressure to the mylohyoid, increasing the proximity between the therapist's hand and Jayden's face, and ultimately providing "air prompts" (Hayden, 2004; L. Slim, personal communication, 2022). However, the fading was very difficult because Jayden often emitted errors without the physical prompt or waited for the therapist to provide the physical prompt. A flexible prompt fading approach, was used in response to Jayden's needs (see Leaf et al., 2022). If Jayden did not imitate correctly, the therapist said, "No," and provided the least assistive prompt to evoke the correct response. These prompts included, verbal feedback (e.g., "Your tongue needs to go up"), gesture (i.e., point towards the mouth), and minimal physical prompts, such as light pressure from the thumb placed anteriorly on the chin, and "air prompts" (see Hayden, 2004; L. Slim, personal communication, 2022; Leaf et al., 2022). In later sessions, the sound was modeled once more after the prompts, to evoke an independent response.

Data for Jayden's imitation of echoic probes is displayed in Figure 6. Jayden never imitated /b/ taught in the vocal imitation training condition. After 55 treatment sessions Jayden correctly imitated both echoic probes for /n/. However, this performance was not sustained. Over the next nine treatment sessions, he correctly imitated both probes on only two occasions. Although Jayden exhibited signs of prompt dependency, he did demonstrate independent production of /n/ during treatment. With the continuous schedule of reinforcement used in treatment, it was suspected that these new responses were easily extinguished in the echoic probes due to the absence of reinforcing stimuli. Therefore, just prior to the presession probes in sessions 66, 67, and 68, Jayden selected preferred stimuli (i.e., tangible fidget toys), which would be used to reinforce correct responses during the echoic probes. These items were held by the therapist, the target vocalization was modeled (e.g., "Say, /n/"), and contingent on a correct response, praise and access to the tangibles were immediately provided. Incorrect responses produced no programmed consequences. That is, if Jayden did not imitate the model, no praise, reinforcement, prompts, or correction procedures were applied. Performance stabilized with this change, and Jayden correctly imitated both echoic probes for /n/ over three consecutive sessions. Jayden became ill after session 68 and was absent for approximately 5 days. Treatment was thus discontinued at this time, and the maintenance probes were conducted 7 days after the last session. Reinforcing stimuli were again selected and available contingent on correct responding during the maintenance probes. Jayden correctly imitated both probes of /n/. Due to his illness, the mastery criterion for Jayden was reduced to correct imitation of two echoic probes across three consecutive sessions. Overall, the total training time to reach this criterion for /n/ was 1,380 trials over 69 treatment sessions.

Figure 6

Echoic Probes for Jayden



Note. The number of correct imitations across two presession echoic probes. B = Baseline; M = Maintenance; SR+ = Praise and tangible reinforcement were delivered contingent on correct imitation

Benjamin

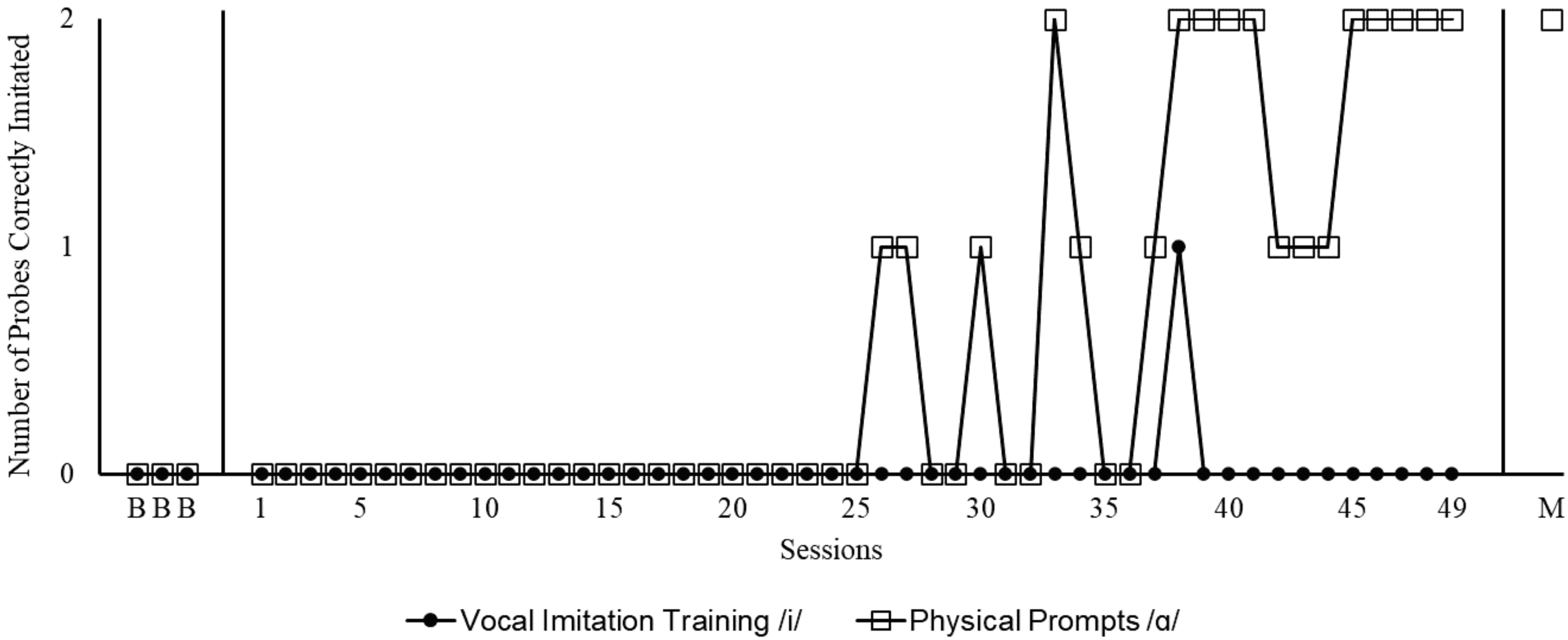
Benjamin's target vocalization for the vocal imitation training condition was /i/. When the therapist modeled the sound, Benjamin typically did not respond. Therefore, on the sixth treatment session, a shaping procedure was implemented. Correct jaw and labial position, even without vocalization, were reinforced. That is, a high jaw position with lip retraction, in the absence of crying, was followed by praise and access to a preferred item or edible. Benjamin inconsistently displayed this articulatory position during treatment, so the vocal requirement was never added.

The target vocalization for the physical prompts condition was /ɑ/. Again, when the therapist modeled the sound, Benjamin typically did not respond. Initially, the therapist emitted the sound, /ɑ/, while providing the physical prompt. As in the vocal imitation training condition, a shaping procedure was implemented on session 6. Praise and access to a preferred edible/tangible were provided for demonstrating the appropriate jaw position, without vocalization, and with the physical prompt. After Benjamin learned to open his mouth in response to the modeled sound and physical prompt, the final component was added to evoke voicing (i.e., second finger placed on the side of the larynx). The therapist modeled the sound, immediately provided all components of the physical prompt while simultaneously voicing, and paused before removing her hands. Any vocalization that occurred with this physical prompt was reinforced and later only those productions that were acoustically similar to the model were reinforced. Physical prompts were faded using a constant time delay of 2 s. During treatment, physical prompts were still used to fine tune production. For example, Benjamin was observed to alter jaw height, producing more of an /ʌ/ vowel sound. By physical providing prompts to help reposition the jaw, he could emit the target vocalization /ɑ/. In the final five treatment sessions, the schedule of reinforcement was reduced to a variable ratio 2, to help strengthen and stabilize responding across the echoic probes.

Data for Benjamin's imitation of the echoic probes is presented in Figure 7. During one presession probe on session 38, Benjamin correctly imitated /i/ on one echoic probe. However, this never occurred again. Correct imitation of /ɑ/ began after treatment session 25 and fluctuated over the next 15 presession probes before stabilizing on sessions 38–41. Correct imitation temporarily decreased over the next three sessions and finally stabilized with correct imitation of both probes across five consecutive sessions. Total training time to reach the mastery criterion for /ɑ/ was 960 trials over 48 treatment sessions. Treatment was discontinued and maintenance probes were conducted 8 days later. Benjamin maintained imitation of /ɑ/ with correct imitation across both echoic probes.

Figure 7

Echoic Probes for Benjamin



Note. The number of correct imitations across two presession echoic probes. B = Baseline; M = Maintenance

Theo

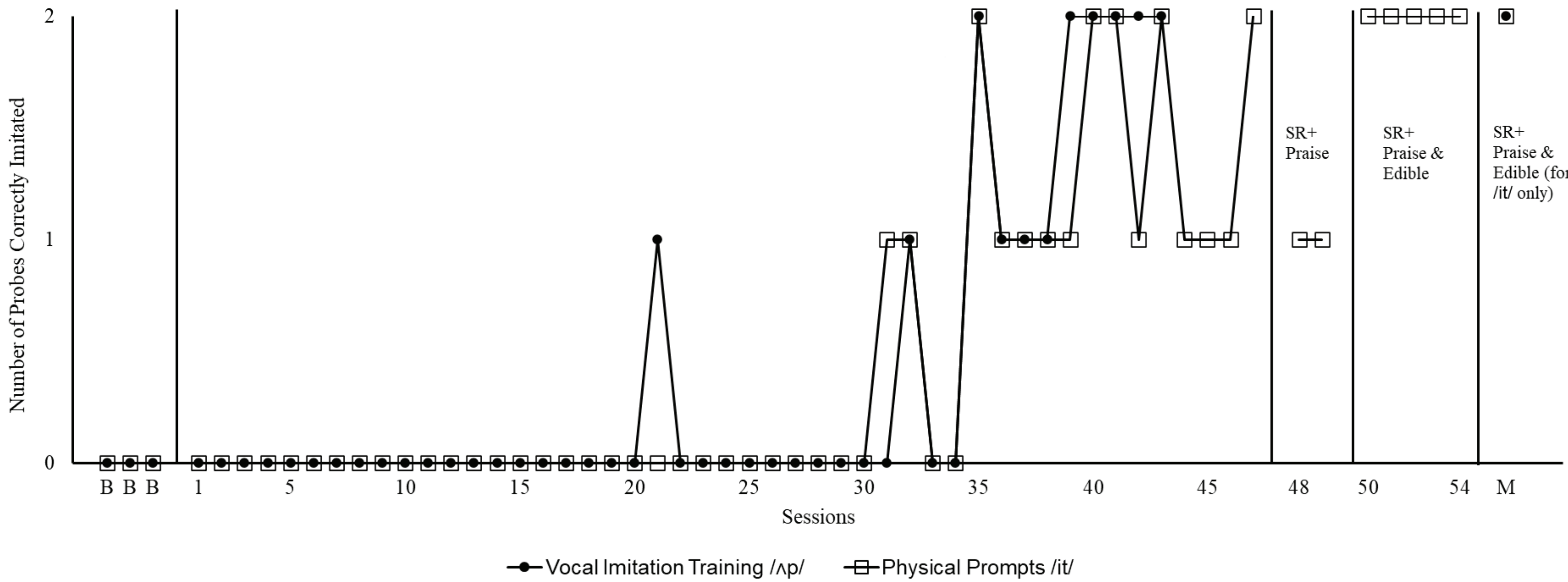
Theo's target vocalization for the vocal imitation training condition was /ʌp/. Errored responses were typically only /ʌ/ or similar vowel sound. Treatment began with imitation of /ʌ/ in isolation followed by /p/. After learning to correctly imitate these sounds in isolation, the full vocal target /ʌp/ was modeled, but with an interresponse time of approximately 2 s. The therapist modeled /ʌ/ and Theo imitated, then the therapist modeled /p/, and Theo again imitated. In subsequent sessions a simultaneous vocal prompt was provided. The therapist modeled the vocal target /ʌp/ then the first phoneme of the syllable, /ʌ/, was slowly presented again, and Theo produced the sound in unison with the therapist. The second phoneme /p/ was presented in the same way and Theo produced the sound in unison with the therapist. The therapist gradually reduced the volume of her vocalization, using only silent articulatory movements, and finally no prompts. Again, most commonly, errors were omission of the final consonant. A modified no–no prompt was used to correct errors. When Theo incorrectly imitated /ʌp/, the therapist said, "No," represented the trial with an emphasis on production of /p/, and simultaneously vocalized the syllable with Theo. Because he was correctly imitating the syllable during treatment, but not in the presession probes, the reinforcement schedule was reduced to a variable ratio 2 in an effort to strengthen and stabilize responding.

The target vocalization for the physical prompts condition was /it/. Errored responses were typically production of only /i/ or similar vowel sound. Treatment began with correct imitation of /i/ in isolation. The therapist modeled the sound, immediately provided the physical prompt, and simultaneously vocalized the sound. After learning to correctly imitate /i/, treatment addressed imitation of /t/ following the same procedural sequences. Next, the two sounds were combined. The full vocal target /it/ was modeled, physical prompts for each phoneme were provided, and the therapist simultaneously vocalized the syllable. Physical prompts were faded in a forward chaining manner by first withholding the prompt for /i/ and later the prompt for /t/. Theo's common error pattern was final consonant deletion; thus, he almost always correctly initiated the response (i.e., produced /i/). When Theo failed to produce the full syllable, the therapist said, "No," provided the physical prompt for /t/ or both phonemes, while simultaneously vocalizing the sounds, and then Theo independently imitated. The physical prompt for /i/ was not necessary to evoke production of the syllable, but it was provided to teach the full sequence. Because Theo was correctly imitating the syllable during treatment, but not in the presession echoic probes, the reinforcement schedule was reduced to a variable ratio 2 once again, in an effort to strengthen and stabilize responding.

Data for Theo's imitation of echoic probes is displayed in Figure 8. Theo acquired imitation of /ʌp/, taught in the vocal imitation training condition before mastering /it/ in the physical prompts condition. After 20 treatment sessions, Theo correctly imitated /ʌp/ on one echoic probe. Performance decreased thereafter with some fluctuation. On session 43, Theo demonstrated mastery of the echoic target /ʌp/, with correct imitation on two probes for five consecutive sessions. Total training time to reach the mastery criterion for /ʌp/ was 840 trials over 42 treatment sessions. Treatment was discontinued, and 8 days later, correct imitation was maintained across two echoic probes. Correct imitation of /it/ first occurred after 30 treatment sessions. Performance on the echoic probes fluctuated for the next 16 sessions. It was suspected that these new responses were easily extinguished in the echoic probes due to the continuous schedule of reinforcement in the treatment sessions. Therefore, praise as determined via parental report to be preferred, was provided contingent on correct imitation in the presession echoic probes on sessions 48 and 49. Yet, responding was not sustained across the two probes. On sessions 50–54 Theo selected a preferred stimulus (i.e., lollipop) prior to the presession echoic probes. The lollipop was held by the therapist, the target vocalization was modeled (e.g., "Say, /it/"), and praise and an edible were immediately provided contingent on a correct response. Incorrect responses produced no programmed consequences. That is, if Theo incorrectly imitated the model (e.g., /i/), no praise, reinforcement, prompts, or correction procedures were delivered. Performance stabilized with this change and Theo correctly imitated both echoic probes for /it/ over five consecutive sessions. Total training time to reach the mastery criterion for /it/ was 1,060 trials over 53 treatment sessions. Maintenance probes were conducted 7 days later, and Theo correctly imitated /it/ on both echoic probes.

Figure 8

Echoic Probes for Theo



Note. The number of correct imitations across two presession echoic probes. B = Baseline; M = Maintenance; SR+ = Praise and/or edible reinforcement were delivered contingent on correct imitation

Interobserver Agreement and Treatment Integrity

The primary experimenter collected data for all sessions. A second observer reviewed videos for at least 33% of echoic probes and treatment sessions for each participant. For the echoic probes, data were collected on correct or incorrect imitation of the modeled sounds. The total number of probes in agreement were divided by the total number of probes and multiplied by 100. Interobserver agreement was 99%, 94%, 100%, and 97% for Sarah, Jayden, Benjamin, and Theo, respectively. In the treatment sessions, data were collected on whether the therapist modeled the target sound/syllable, provided the prompt if necessary, and delivered the reinforcing stimuli for correct responses. In addition, data were collected on whether the therapist offered praise for other appropriate behaviors and a break away from the table. The total number of trials correctly implemented was divided by the total number of trials and multiplied by 100. Treatment integrity was 99%, 100%, 99%, and 99% for Sarah, Jayden, Benjamin, and Theo, respectively.

Social Validity

The questionnaire items and results are displayed in Table 2. Generally, agreement for all questionnaire items was high (i.e., *somewhat agree* to *strongly agree*). Each respondent always had the same level of agreement for the two treatments or higher agreement for the physical prompts treatment. An exception was Sarah's mother, who strongly agreed that the vocal imitation training was effective and would recommend it for other children, but only somewhat agreed with these questionnaire items regarding the physical prompts treatment. According to the comments, several respondents demonstrated a preference for the physical prompts treatment. For example, after reviewing the video of the physical prompts treatment, Theo's father commented "Huge improvement!" Likewise, Theo's behavior technician remarked that the physical prompts produced a noticeable change and were not aversive to him. Even though she found the vocal imitation training to be effective, she felt that Theo had greater difficulty attending. As another example, Jayden's behavior technician said, "It was so amazing to see the progress he made throughout this study!" Jayden's mother also agreed that the physical prompts helped him learn to produce sounds by teaching him to position his mouth accordingly. Then again, based on her past experience, she felt that the physical prompts were not consistently effective for Jayden and may not be suitable for all children because many may feel uncomfortable with touch to their face. Acknowledging that the vocal imitation training is likely effective for many children, she explained that it had been unsuccessful for Jayden. Finally, Benjamin's behavior technician commented that the physical prompts were both effective and efficient for teaching echoic behavior and that Benjamin seemed comfortable and content with the procedures. Moreover, she stated that Benjamin, "demonstrated a faster than typical skill acquisition rate when compared to other skills currently in intervention". In short, parents and behavior technicians found both the vocal imitation training and the physical prompts to be acceptable and effective treatments, although slight differences in Likert scores and feedback in the comments, indicated an increased preference for the physical prompts treatment.

Table 2

Mean and Standard Deviations of Scores on the Social Validity Questionnaire

Questionnaire item	Behavior technicians (*n* = 4)		Parents (*n* = 4)	
	Vocal imitation training	Physical prompts	Vocal imitation training	Physical prompts
Treatment specific				
I found this treatment to be acceptable.	4.8 (0.5)	5.0 (0.0)	5.0 (0.0)	5.0 (0.0)
I found this treatment to be effective.	4.3 (0.5)	4.8 (0.6)	4.3 (1.5)	4.5 (0.6)
I would recommend this treatment for other clients/children.	4.5 (0.5)	5.0 (0.0)	4.5 (1.0)	4.3 (1.0)
Overall				
The intended outcome of these interventions – to improve my child/client's imitation of vocal sounds – is important to me.	5.0 (0.0)		5.0 (0.0)	
I have noticed an improvement in my child/client's vocal imitation following participation in these treatments.	4.3 (1.0)		4.0 (1.2)	
I feel comfortable with the procedures that were used with my child/client.	5.0 (0.0)		5.0 (0.0)	
I have noticed an improvement in my child/client's spontaneous vocalizations following participation in these interventions.	4.3 (1.5)		4.8 (0.5)	
I have heard my child/client emit new sounds as a result of these interventions.	4.0 (1.4)		4.0 (1.2)	
Overall, participating in these treatments, made a difference in my child/client's vocal abilities.	4.5 (1.0)		4.8 (0.5)	

Note. Standard deviations are presented in parentheses. Scores were based on a 5–point Likert scale of agreement. 1 = strongly disagree, 2 = somewhat disagree, 3 = neither agree nor disagree, 4 = somewhat agree, 5 = strongly agree.

Discussion

The purpose of the current study was to compare the effectiveness of two treatment techniques, vocal imitation training and physical prompts, on echoic responses in children on the autism spectrum who had limited vocal behavior and/or severe speech sound disorders. All participants acquired new echoic responses following treatment. These findings contribute to the existing echoic literature by demonstrating the effectiveness of imitation (e.g., Carroll & Klatt, 2008; Civindi-Motta et al., 2017; Dyer et al., 1987; Esch et al., 2005; Koegel et al., 1988; Tarbox et al., 2009), and physical prompts (e.g., Aravamudhan & Awasthi, 2020, 2021; Lovaas et al, 1966;) on the development of echoic responses.

One of the unique aspects of the present study is the within–subject comparison of methods. When comparing the effectiveness of the instructional methods, differences were observed across participants. Interestingly, Sarah and Theo acquired echoics under both treatment conditions, but the vocal target taught using vocal imitation training was mastered in fewer sessions than the target taught using physical prompts. Yet Jayden and Benjamin, who presented with more limited vocal communication and imitation at the start of the study, only acquired the echoic responses taught using physical prompts. These two participants did not acquire sounds in the vocal imitation training condition. All participants maintained the newly trained echoic responses when treatment was discontinued, except for Sarah, who only maintained the target taught in the vocal imitation training condition. Overall, these results suggest that there may have been an advantage with the physical prompts for learners with limited echoic repertoires. This is a very interesting finding that warrants replication. It may be that physical prompts are especially relevant for learners with minimal vocal behavior and/or certain speech sound disorders.

Motor speech disorders, especially apraxia, are often believed to be the underlying cause of the vocal communication deficits in minimally vocal-verbal children on the autism spectrum (Adams, 1998; Beiting, 2022; Chenausky et al., 2019; Gernsbacher et al., 2008; Shriberg, 2010; Tierney et al., 2015). Even so, an accurate diagnosis is challenging in this population. Current evaluation procedures and diagnostic criteria are based on the patterns and characteristics of vocal behavior, and limited vocal output, perhaps ironically, prevents a thorough evaluation of motor speech function (Chenausky et al., 2019). The nature of the vocal communication deficits exhibited by the participants in the current study is unknown. Even still, the physical prompts were effective for all participants, and it was the only procedure that successfully established echoic responses with Jayden and Benjamin. Although it is impossible to conclude, it may be that all participants presented with the same core deficits (e.g., apraxia of speech, speech delay) and the physical prompts effectively addressed these deficits. Or, perhaps, physical prompts are a sort of broad-spectrum treatment, which can effectively be applied to a variety of speech sound disorders. Either way, additional research is needed to examine the effects of physical prompts as it relates to participant characteristics and the nature of the disorder.

Sarah and Theo did, indeed, learn to imitate the target taught with physical prompts, but they acquired the target in the vocal imitation training condition in fewer trials. One possible reason for this increased rate of acquisition is a proximal history with the prompt procedures used in the vocal imitation training condition. A few studies have examined the effects of previous exposure to a certain prompt type, on the acquisition rate of intraverbal behavior (i.e., Coon & Miguel, 2012; Kay et al., 2020; Roncati et al., 2019). More specifically, these studies compared the efficiency of echoic and tact prompts following exposure training. Both prompt types were effective, but all studies found that participants acquired new intraverbal responses in fewer teaching trials with the prompt procedure most recently used. According to Coon and Miguel (2012), the proximal history of reinforcement establishes a stronger degree of stimulus control and increases the rate of acquisition. It is unknown if the participants in the present study were recently exposed to the same prompt procedures used in treatment. However, vocal imitation training is a relatively common approach to teaching echoic responses and improving speech intelligibility in children on the autism spectrum (see Beiting, 2022), and it is possible that these procedures were used to address other vocal behaviors. If Sarah and Theo were exposed to these same prompt types, it might explain why the vocal imitation training was more efficient. Future research might evaluate prompt types and prompt systems prior to intervention (Kay et al., 2020; see also Seaver & Bourret, 2014). Moreover, it might be worthwhile to examine proximal history with certain prompt types and the effects on vocal behavior. Then, we may even be able to predict efficiency of certain methods based on a participant's previous exposure (Roncati et al., 2019).

In the present study, only one vocal target was taught under each treatment condition, but some studies have reported an increased rate of acquisition as echoic behaviors were learned. Lovaas et al. (1966) found that as additional targets were trained, the number of days and trials to reach mastery decreased, with many responses being acquired the day they were introduced. Tarbox et al. (2009) also found that for each of three participants, the initial vocal target required relatively more treatment time, but subsequent targets were acquired more efficiently. Thus, simply establishing a history of reinforcement with certain methods, even prior to intervention, may improve the rate of acquisition (Coon & Miguel, 2012; Roncati et al., 2019). Future examinations may evaluate echoic behavior across targets and measure efficiency through a multiple baseline across behaviors design (e.g., Tarbox et al., 2009). Generalization probes could be conducted intermittently to evaluate changes to untrained targets (Aravamudhan & Awasthi, 2020).

Examining variables that affect efficiency of acquisition will be an important area for future research. Although the physical prompts were effective for all participants, the average number of trials to mastery was 1,105 (range 960–1,380) across an average of 54.8 treatment sessions (range 48–67). In the vocal imitation training, which was only effective for two of the participants, the average number of trials to mastery was 870 (range 840–900) across an average of 43.5 treatment sessions (range 42–45). These ranges are similar to those found in previous research. Based on the data reported by Lovaas et al. (1966), mastery of the initial echoic response with physical prompts took approximately 23,760 trials. As they explained, treatment was conducted 6 days a week, 7 hours a day, with a 15-minute break every hour. Vocal targets were modeled on average, every 10 seconds, and the first vocal target, “baby,” was acquired on day 12 of treatment. As another example, Aravamudhan and Awasthi (2020) reported that Participant A took on average 532 trials (range 72–1230), within 54.75 sessions (range 6–125), over 9.75 days (range 1–23) to master the target words in training. Participant B, whose training included a physical prompt for one of the targets, took on average 730 trials (range 264–2044), within 77.11 sessions (range 30–201), over 13.88 days (range 5–36). Rates also varied across specific targets. Each of their participants varied by as much as 1,000+ trials across targets, to reach the mastery criterion. Similarly, Dyer et al. (1987) found, that for their participant, the number of trials varied from 30–220 for each of the target phonemes. Perhaps then, there are certain idiosyncratic features of speech sounds/syllables that make some inherently easier to produce than others.

Despite attempts to match the vocal targets for each participant in the present study with regard to developmental norms and phonemic characteristics, one target may have required less effort to produce based on the child's speech motor coordination and control. For example, Theo's vocal targets /ʌp/ and /it/ are both VC syllables with voiceless stop consonants. During preassessment, Theo correctly produced all phonemes in other contexts. The vocal target /ʌp/, which was taught in the vocal imitation training condition, was acquired in fewer trials than /it/. When compared to /t/, earlier ages of acquisition have been reported for /p/ (Roth & Worthington, 2018), and so /ʌp/ may be cognitively or motorically simpler to produce. Also, as a bilabial stop, the articulatory movements necessary for /p/ are more transparent than /t/ when modeled, possibly making imitation easier. Jayden's vocal targets /b/ and /n/ share similar differences. That is, some studies have found earlier ages of acquisition for /b/ (Roth & Worthington, 2018), and, as a bilabial stop, the articulatory movements for production of /b/ are more transparent than for the production of /n/. However, Jayden only acquired /n/, which was taught in the physical prompts condition, and he never imitated /b/ under the vocal imitation training condition.

Benjamin's vocal targets were /i/, addressed in vocal imitation training, and /ɑ/ taught with physical prompts. He never mastered imitation of /i/, but learned to imitate /ɑ/, and he maintained imitation when treatment was discontinued. According to the PROMPT methodology and the Motor Speech Hierarchy, the stages of speech motor control and coordination are hierarchically dependent, such that earlier developing stages directly influence later developing stages (Hayden, 2004, 2006, 2021). As an open vowel, Benjamin's target /ɑ/ required mandibular movement in the vertical plane (i.e., Stage 3 of the Motor Speech Hierarchy), whereas /i/, with labial retraction and minimal jaw opening, occurred in a horizontal plane of movement (i.e., Stage 4 of the Motor Speech Hierarchy; Hayden, 2006). Consequently, treatment of speech motor control for vowel production should initially focus on phonation through graded jaw movement. Beginning with sounds like /ɑ/, which requires a fully open mouth, treatment then progresses to sounds like /i/, with more restricted aperture (Hayden, 2004). So, it is possible, relative to his speech motor function and deficits across subsystems, that the labial movements for /i/ were more complex for Benjamin, making /ɑ/ a simpler, more attainable sound.

Sarah learned to imitate both /fʌni/ and /muvi/. However, contrary to the progressive, systematic design of the PROMPT framework, she demonstrated better performance with more intricate phonemes and syllables typically targeted in later phases of treatment (see Hayden 2004, 2006, 2021). Voiced sounds require less motor control and should be targeted before unvoiced sounds (Hayden, 2004). Despite this, Sarah acquired /fʌni/, which requires a higher level of voicing control to transition from a voiceless consonant to a vowel, in fewer trials than /muvi/, which simply requires sustained phonation. Furthermore, articulatory movements within a single plane (i.e., Phase 1) are targeted before teaching coordination across multiple planes of movement (i.e., Phase 2; Hayden, 2004). The vocal target /muvi/ occurred only in the horizontal plane, whereas /fʌni/ crossed planes of movement. Beginning with /fʌ/ in the vertical plane, it ends with /ni/ in the horizontal plane. Theoretically, /muvi/ may require less speech motor coordination and control than /fʌni/, making it inherently easier to imitate. Yet, this was not reflected in Sarah's performance, as she required more training trials to imitate /muvi/ and did not maintain imitation once treatment was discontinued.

One possible idiosyncratic feature that may have influenced the acquisition of specific sounds/syllables is each participant's conditioning and reinforcement history with vocal stimuli. Prior conditioning of speech sounds, and/or a history of reinforcement, could not be adequately assessed or controlled, and likely varied across participants. Young children hear speech sounds produced by their caregivers, and while these auditory stimuli are initially neutral, they become conditioned reinforcers through their temporal association with reinforcing stimuli (Carrol & Klatt, 2008; Esch et al., 2005; Skinner, 1957). As Esch et al. (2005) explained, children then begin approximating the speech sounds they hear, and the articulatory movements of sound production are automatically reinforced by the auditory response product generated by those movements. Consequently, through the social feedback loop, caregivers respond to these speech–like vocalizations, which helps shape, expand, and strengthen the child's vocal behavior (Warlaumont et al., 2014). The conditioning and reinforcement history, for each of the participants in the current study, is unknown. It is, however, possible that although the vocal targets were not under echoic control at the start of the study, they may have been previously conditioned and reinforced in other contexts. Such a history, especially for Jayden or Benjamin, may have offered an advantage when learning to imitate. The sensory consequences (i.e., feeling and hearing the sound) may have had a reinforcing effect on the articulatory movements, and production of the sound. Emitting the vocal target(s) within noncontextual, spontaneous vocal play/babbling may have facilitated the motor coordination and control necessary for production; it also could have evoked social attention. Then, in the current study, presentation of the conditioned auditory stimulus, as opposed to a neutral stimulus, might have increased motivation, attention to the model, and initiation of articulatory movements, essentially making the sound easier to acquire. In other words, it is unclear if the vocal targets acquired in the current study were new sounds/syllables or existing vocalizations brought under control of new stimulus conditions. However, even if such a history of conditioning and reinforcement existed, it may have had little impact on performance. Previous investigations have attempted to condition vocal sounds and consequently bring them under echoic control with differential reinforcement, yet the results have been mixed (e.g., Carroll & Klatt, 2008; Esch et al., 2005).

A limitation of the current study is that direct preassessment measures only examined correct/incorrect imitation of phonemes in isolation, and across progressively complex syllable shapes. But vocal speech is a multiply controlled, complex behavior which integrates auditory processing, cognition, motor planning/programming, and execution of fine-tuned anatomical movements (Hoff, 2001; Terband et al., 2019). Such behaviors are all under the influence of environmental variables and maintained by the verbal community (Skinner, 1957). More comprehensive evaluations of vocal behavior should include an oral mechanism exam, assessment of speech motor function (e.g., *Verbal Motor Production Assessment for Children*, probing sounds/words according to the Motor Speech Hierarchy), evaluation of vocalizations across other verbal operants, and an analysis of noncontextual, spontaneous speech. According to Beiting (2022), the most effective and efficient treatments are based on the specific nature of the speech sound disorder. For example, motor speech disorders, such as apraxia and dysarthria warrant treatments that address motor coordination, whereas phonological disorders require more linguistically based treatment. Unfortunately, the exact reasons many children on the autism spectrum fail to acquire functional vocal speech is unknown (Beiting, 2022; Rogers et al., 2006; Tager–Flushberg & Kasari, 2013). Differential diagnosis of speech sound disorders in autism is challenging (Beiting, 2022), but will be important to the study of vocal communication and the development of treatments that align with the underlying deficits. By evaluating speech productions across multiple contexts, the core deficits, topography of current responses, controlling variables, and reinforcing properties of auditory stimuli can potentially be identified, therefore improving target selection and treatment implementation.

Vocal communication and speech sound disorders in minimally vocal-verbal individuals is a relatively understudied area of autism (Beiting, 2022; DeSouza et al., 2017; Koegel et al., 2020). This presents unique opportunities, in both research and clinical practice, for speech-language pathologists and behavior analysts to combine efforts and develop efficacious interventions. Speech-language pathologists have extensive knowledge of both typical and atypical patterns of human development as it relates to vocal speech, including anatomy, physiology, respiration, oral-motor coordination, articulation, phonation, and resonance (Council on Academic Accreditation in Audiology and Speech-Language Pathology, 2023). Their training covers the cognitive, social, and cultural aspects of vocal communication, as well as the etiology and characteristics of various disorders and differences. Hence, their practice commonly includes assessment, diagnosis, and treatment of vocal communication deficits (Council on Academic Accreditation in Audiology and Speech-Language Pathology, 2023; LaFrance et al., 2019). With this background and expertise, speech-language pathologists possess the foundational skills necessary for additional training in treatment approaches such as PROMPT, which emphasizes thorough assessment of motor speech function and precise treatment of articulatory movements (Hayden 2004, 2006, 2021). Behavior analysts, on the other hand, have a strong background in the science of human behavior. Specifically, their knowledge includes the concepts, principles, and procedures of behavior analysis and how these may be applied in a clinical context to improve the human condition (Behavior Analyst Certification Board, n.d.). In the treatment of vocal communication deficits, behavior analysts can establish contingencies of reinforcement to increase vocal responses and bring them under control of relevant stimuli. By determining the appropriate schedule of reinforcement, they can strengthen new responses and then make adjustments that maintain consistent, stable responding over time. When designing a new treatment program, behavior analysts are able to assess other areas important to performance, such as stimulus preferences and effectiveness of prompts. They can also help program for generalization, determine ideal teaching arrangements (e.g., discrete trial instruction, naturalistic teaching), and recommend additional behavior change procedures that are relevant to the treatment of vocal communication. The behavior analysts' knowledge of single-subject experimental design is especially useful in evaluating the effects of a treatment procedure implemented in an applied setting. As such, they can assist with measuring and analyzing changes in vocal behavior according to various dimensions such as frequency, rate, percentage, and topography. Uniting the expertise of speech-language pathologists and behavior analysts is likely to generate more comprehensive, effective, and efficient treatments than either discipline could accomplish independently.

With their capacity for treating vocal communication, speech-language pathologists and behavior analysts may also need to collaboratively address learning readiness skills where needed. Vocal communication deficits are often one of the earliest, most pressing concerns of parents of children on the autism spectrum (Mayo et al., 2013). Still, many young children lack the necessary prerequisite skills to fully participate and benefit from intervention. Chen et al., (2019) reportedly excluded five children from participating in their research because they did not engage with treatment materials or did not cooperate with experimenters. Chenausky et al., (2016) excluded as many as 30 children for various reasons, one of which was the inability to participate in table-top activities for at least 15 min. In the current study, two children were excluded because adult-directed demands resulted in acts of protest or aggression. Speech-language pathologists and behavior analysts can help young learners acquire important learning readiness skills such as eye contact, sitting, attending to an activity, cooperating with basic prompts and instructions, and understanding basic contingencies (e.g., first I do, then I get) as efficiently as possible, so that they may receive appropriate and effective treatment (Taylor & Fisher, 2010). Otherwise, vocal communication development may be further hindered.

Such a partnership between speech-language pathology and behavior analysis has previously generated evidence-based treatment packages that have successfully improved communication in children with developmental disabilities (Koenig & Gerenser, 2006). The *Picture Exchange Communication System* is an alternative communication system developed by a behavior analyst and a speech-language pathologist (Bondy & Frost, 2001). Through this system, children with limited vocal-verbal abilities have successfully learned to initiate interactions, emit mands, and in some cases, produce vocal speech and exhibit more advanced language (Preston & Carter, 2009). As another example, *Pivotal Response Training*, which is a combination of elements taken from the Natural Language Paradigm and applied behavior analysis, has improved self-initiations, language, play skills, and sustained interactions in young children (Verschuur et al., 2013). Incorporating physical prompts, to address speech motor sound production, into empirically validated interventions such as these would be an interesting area of future research between speech-language pathology and applied behavior analysis. For example, while exchanging pictures to request preferred items/activities, according to the Picture Exchange Communication System, physical prompts with simultaneous vocal prompting could be provided to evoke speech production. After reinforcing vocal approximations in Pivotal Response Training, the instructor might implement physical prompts to encourage responses that more closely match the adult model. Other opportunities for collaborative research include auditory match-to-sample procedures and precision teaching. Through discrimination training, auditory match-to-sample procedures have effectively improved echoic responses in children on the autism spectrum with limited vocal imitation (Brown, 2005; Choi et al., 2015; Du et al., 2017; Speckman–Collins et al., 2007). With the addition of physical prompts, children may learn the articulatory movements more easily, further improving efficiency, and facilitating a generalized echoic repertoire. Combining physical prompts with the principles of precision teaching might yield the highest rates of acquisition. Aravamudhan and Awasthi (2021) used precision teaching and a physical prompt to address imitation of /kʌ/. After only 16 days and a total of 24 min of treatment, the target was mastered. Since fluency is emphasized in precision teaching, this combination would potentially improve overall speech intelligibility. By building on the existing science, speech-language pathology and applied behavior analysis may come together to further improve the effectiveness and efficiency of these interventions.

Social validity was assessed via a subjective questionnaire completed by behavior technicians and parents of participants. In general, agreement for all questionnaire items was high (i.e., *somewhat agree* to *strongly agree*) indicating social validation of the overall goal and treatment procedures. Unanimously, respondents felt comfortable with the procedures and strongly agreed that the intended outcome, to improve vocal behavior, was important. Future studies might evaluate other aspects of treatment acceptability and the importance of outcomes such as efficiency, cost-effectiveness, feasibility, and the amount of meaningful change (see Foster & Mash, 1999). Moreover, social acceptability data should be collected from other relevant consumers including recipients of the treatment, potential clients, speech-language pathologists, and behavior analysts.

The current study utilized discrete-trial instruction, which is well established for the treatment of speech sound disorders, in conjunction with other behavioral techniques (Hedge and Pena-Brooks, 2008). The procedures were implemented according to a “structured, yet flexible approach" as described by Leaf et al. (2016, p. 721), whereby treatment is customized, continuously analyzed, and modified, according to the needs and responses of the client. Despite efforts to thoroughly describe the procedures, replication may be difficult. Instructions for each trial were varied to offer a more naturalized presentation (e.g., "Look at me, /ɑ/," "Ready? /fʌ/," "Say, /ʌp/," "/n/"), although this created inconsistency across trials (Leaf et al., 2016). Prompts included simultaneous voicing, verbal feedback and cueing (e.g., "Use your lips," "Your tongue needs to go up"), as well as gestures (i.e., pointing toward face), and physical prompts (in the physical prompts condition). Consistent with a flexible, progressive approach to treatment, these prompts were delivered and faded in a dynamic manner. Prompt fading, depended on several factors, including the child's attention, recent response, the next sound in the sequence, the critical feature missing from production (e.g., tongue tip elevation for /n/, bilabial closure for /m/), and common error patterns. (Leaf et al., 2016, 2022). That is, decisions were often made "in the moment" and "on a case-by-case basis" (Leaf et al., 2016, p. 726). Future research should evaluate various prompt types (e.g., verbal, physical, within-stimulus) and prompt systems (e.g., use of constant time delay, most-to-least) with clear, hierarchical fading guidelines that effectively establish and maintain echoics in learners with minimal vocal behavior.

Implementing the physical prompts from PROMPT requires precise touch, pressure, timing, and movement to evoke speech, and facilitate coarticulation (Hayden, 2004). Consistent application is difficult, as participants have different anatomical structures, variations in movement, and unique error patterns. Determining the appropriate prompt type (e.g., parameter, surface, complex, or syllable) also requires consideration of several factors that may vary depending on the client's needs and context. When Sarah was learning to sequence all phonemes in the target /muvi/, the parameter prompt for /m/ was applied, but the surface prompt may have been more appropriate given the need to focus on timing and sequencing (Hayden, 2004). Due to the stimulus control exerted by the parameter prompt, it was however continued even at the word level. In short, physical prompts for vocal behavior require a certain level of behavioral artistry, careful analysis, and professional judgment.

Conventional vocal imitation training typically includes only a model and differential reinforcement (Civindi-Motta et al., 2017). In addition to the various prompts provided, a shaping procedure was added for three of the participants in the current study. Because of their observed errors, it was added to both treatment conditions within the first week. It proved effective for Sarah and Theo in both conditions, and Benjamin only in the physical prompts condition. Benjamin was taught to first open his mouth in response to the vocal model and physical prompts and subsequently to vocalize /ɑ/. Initially, Theo's treatment taught the individual phonemes of each syllable and later focused on combining them. Sarah first learned to imitate individual phonemes, then sequences of sounds, and finally the entire word. Whether these procedures actually constitute shaping, or are more correctly described as chaining, is a conceptual question (see Tarbox et al., 2009). Each link in the chain (e.g., each phoneme, each articulatory movement) was directly taught and then linked together to create a syllable structure, as is typical of chaining. Yet successive approximations (e.g., opening mouth, sound in isolation) toward a terminal behavior (e.g., imitating /ɑ/, producing syllable), as is characteristic of shaping, were also included. Nonetheless, these prompt and behavior change procedures, were successfully incorporated into the discrete-trial instruction. However, more naturalistic teaching arrangements are also suitable for improving vocal behavior with these procedures and should be experimentally evaluated. They further capitalize on motivation, promote generalization and maintenance of new behavior, and allow for the training of other vocal–verbal operants, such as mands, tacts, and intraverbals (Charlop-Christy et al., 1999).

The mand is believed by many behavior analysts to be the most important verbal operant (Sundberg, 2015). It is, undoubtedly, important in teaching language to children on the autism spectrum. Then again, many behavior analysts identify the intraverbal as the most important because it seems most pervading (Sundberg, 2015). Skinner (1957), perhaps surprisingly, wrote that the tact was the most important verbal operant, due to its nonverbal stimulus control and its benefit to the listener, among other reasons. But it can be argued that the echoic deserves special consideration, especially in teaching vocal behavior to individuals on the autism spectrum who historically, have had difficulty acquiring spoken language.

The echoic may be an important component of cumulative-hierarchical learning, or perhaps best conceptualized as a behavioral cusp, where learning one behavior enables the learning of another behavior (Bosch & Hixon, 2004; Pecukonis et al., 2019). Echoic behavior is apparent in early childhood and is critical to the acquisition of spoken language (Brown, 2005; Hoff, 2001). If we teach children to imitate the vocalizations they hear, and caregivers then respond to these vocalizations, then the vocal behavior may be further increased and expanded through this social feedback (Warlaumont et al., 2014). As Skinner (1957) explained, even a minimal echoic repertoire is useful for evoking responses and bringing them under the control of new stimuli to further expand language. Research has demonstrated successful use of echoic prompts to teach vocal mands (Drash et al., 1999; Jennett et al., 2008), intraverbals (Kodak et al., 2012; Xu et al., 2020), and tacts (Barbera & Kubina, 2005; Bloh, 2008). The echoic is also a critical component of the *name relation*, described by Horne and Lowe (1996), as the basic unit of language whereby children learn to respond as listeners to their own speech. More than just a foundational skill, the echoic is also useful in advanced language, as it allows the listener to follow directions more accurately, facilitates memory/recall, maintains the topic in conversation, and repairs communication breakdowns through repeating, rephrasing, and asking for clarification (see Skinner, 1957). Future studies might investigate the echoic as a behavioral cusp by measuring additional changes in behavior following the acquisition of an echoic repertoire (Bosch & Hixson, 2004).

According to Aravamudhan and Awasthi (2020), vocal behavior should first be taught as an echoic before teaching speech production across other verbal operants. However, common procedures, namely conventional vocal imitation training, are only effective for children with existing vocal behaviors (Chen et al., 2019; Civindi-Motta et al., 2017) and verbal imitation abilities (Biller & Johnson, 2019). This can be seen in the present study where only two participants, those with relatively higher imitation and speech abilities, acquired targets with vocal imitation training. Consequently, children with minimal vocal behavior are often ineligible for echoic training, and the focus of treatment often shifts to augmentative/alternative communication. But, as Lovaas is often remembered for saying, “If a child cannot learn in the way we teach, then we must teach in the way the child can learn” (Leaf et al., 2016). With approximately 25% of children on the autism spectrum exhibiting limited to absent vocal–verbal communication even after intervention, we must teach vocal behavior in a way they will learn (Anderson et al., 2007; Norrelgen et al, 2015; Rose et al., 2016). All participants in the current study acquired the vocal targets in the physical prompts condition. In fact, Jayden and Benjamin, who presented with the most impaired motor imitation and restricted echoic repertoire, only acquired the vocal targets under the physical prompts condition. Overall, the current study provides evidence for the use of physical prompts in effectively teaching echoic behavior to children on the autism spectrum with minimal vocal behavior and/or severe speech sound disorders. This is an intriguing finding that warrants further exploration and further comparison.

Conclusion

Vocal behavior has been associated with higher verbal abilities (Biller & Johnson, 2019; Paul et al., 2008; Saul & Norbury, 2020; Smith et al., 2007; Thurm et al., 2007; Wetherby et al., 2007; Yoder et al., 2015) and improved outcomes in adulthood (Howlin et al., 2004; Venter et al., 1992), not to mention the positive effects it likely has on social relationships and overall quality of life. Yet, the research on echoic behavior and speech sound disorders in minimally vocal-verbal children on the autism spectrum is dearth (Beiting, 2022; DeSouza et al., 2017; Koegel et al., 2020). With the desire for spoken language (Mayo et al., 2013) and the right to the most effective and efficient means of communication (Brady et al., 2016), continued investigation in this area is critical. Based on the findings of the present study, it may be that physical prompts, which worked for all participants, are particularly well suited to individuals with limited imitation and vocal–verbal abilities. If this finding is replicated, it could lead to new, effective methods to establish vocal communication with such learners. The combined expertise of speech-language pathologists and behavior analysts is quite valuable to illuminate this area of research and practice. It is necessary to understand how individual differences may impact acquisition and identify the most effective prompt types and systems to improve efficiency and sustain responding. In the meantime, these findings underscore the importance of continued exploration of this area and highlight the need for continued interprofessional collaboration in identifying best practices to build echoic repertories and establish vocal communication.

References

Adams, L. (1998). Oral–motor and motor–speech characteristics of children with autism. *Focus on Autism and Other Developmental Disabilities*, *13*(2), 108–112. https://doi.org/10.1177/108835769801300207

Adolfsson, J., & Steineck, G. (2000). Prognostic and treatment–predictive factors—Is there a difference? *Prostate Cancer and Prostatic Diseases*, *3*(4), 265–268. https://doi.org/10.1038/sj.pcan.4500490

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Speech–Language–Hearing Association (n.d.–a). *Childhood apraxia of speech (practice portal)*. https://www.asha.org/practice–portal/clinical–topics/childhood–apraxia–of–speech/#collapse_9

American Speech–Language Hearing Association. (n.d.–b). *Dysarthria*. https://www.asha.org/public/speech/disorders/dysarthria/

American Speech–Language Hearing Association. (n.d.–c). *Speech sound disorders – Articulation and phonology*. https://www.asha.org/practice–portal/clinical–topics/articulation–and–phonology/#collapse_1

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K., & Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. https://doi.org/10.1037/0022–006X.75.4.594

Aravamudhan, S., & Awasthi, S. (2020). Behavioral interventions to treat speech sound disorders in children with autism. *Behavior Analysis in Practice*, *13*(1), 174–185. https://doi.org/10.1007/s40617–019–00362–5

Aravamudhan, S., & Awasthi, S. (2021). The use of prompts and precision teaching to address speech sound disorders in a 17–year–old girl with autism. *Behavior Analysis in Practice*, *14*(3), 644–659. https://doi.org/10.1007/s40617–020–00470–7

Barbera, M. L., & Kubina, R. M. (2005). Using transfer procedures to teach tacts to a child with autism. *The Analysis of Verbal Behavior*, *21*(1), 155–161. https://doi.org/10.1007/bf03393017

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*(1), 85–123. https://doi.org/10.1017/s0305000900008680

Bauman–Waengler, J. (2004). *Articulation and phonological impairments: A clinical focus* (2nd ed.). Pearson Education.

Behavior Analyst Certification Board. (n.d.) *About behavior analysis*. https://www.bacb.com/about-behavior-analysis/

Beiting, M. (2022). Diagnosis and treatment of childhood apraxia of speech among children with autism: Narrative review and clinical recommendations. *Language, Speech & Hearing Services in Schools*, *53*(4), 947–968. https://doi.org/10.1044/2022_LSHSS–21–00162

Beiting, M., & Maas, E. (2021). Autism–centered therapy for childhood apraxia of speech (ACT4CAS): A single–case experimental design study. *American Journal of Speech–Language Pathology*, *30*(3), 1525–1541. https://doi.org/10.1044/2020_AJSLP–20–00131

Bernard–Opitz, V., Sriram, N., & Sapuan, S. (1999). Enhancing vocal imitations in children with autism using the IBM Speech Viewer. *Autism*, *3*(2), 131–147. https://doi.org/10.1177/1362361399003002004

Biller, M. F., & Johnson, C. J. (2019). Social–cognitive and speech sound production abilities of minimally verbal children with autism spectrum disorders. *American Journal of Speech–Language Pathology*, *28*(2), 377–393. https://doi.org/10.1044/2018_AJSLP–18–0008

Bloh, C. (2008). Assessing transfer of stimulus control procedures across learners with autism. *The Analysis of Verbal Behavior*, *24*(1), 87–101. https://doi.org/10.1007/bf03393059

Bondy, A., & Frost, L. (2001). The Picture Exchange Communication System. *Behavior Modification*, *25*(5), 725-744. https://doi.org/10.1177/0145445501255004

Bopp, K. D., Mirenda, P., & Zumbo, B. D. (2009). Behavior predictors of language development over 2 years in children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 52*(5), 1106–1120. https://doi.org/10.1044/1092–4388

Bosch, S., & Hixson, M. D. (2004). The final piece to a complete science of behavior: Behavior development and behavioral cusps. *The Behavior Analyst Today*, *5*(3), 244–254. https://doi.org/10.1037/h0100033

Brady, N. C., Bruce, S., Goldman, A., Erickson, K., Mineo, B., Ogletree, B. T., Paul, D., Romski, M., Sevcik, R., Siegel, E., Schoonover, J., Snell, M., Sylvester, L., & Wilkinson, K. (2016). Communication services and supports for individuals with severe disabilities: Guidance for assessment and intervention. *American Journal on Intellectual and Developmental Disabilities, 121*(2), 121–138. https://www.asha.org/njc/communication–bill–of–rights/

Brisson, J., Martel, K., Serres, J., Sirois, S., & Adrien, J. L. (2014). Acoustic analysis of oral productions of infants later diagnosed with autism and their mother. *Infant Mental Health Journal*, *35*(3), 285–295. https://doi.org/10.1002/imhj.21442

Brown, M. C. (2005). The effects of the acquisition of a generalized auditory word match –to – sample repertoire on the echoic repertoire under mand and tact conditions (Publication No. 3159725). [Doctoral dissertation, Columbia University]. ProQuest Dissertations and Theses Global. https://www.proquest.com/dissertations–theses/effects–acquisition–generalized–auditory–word/docview/305016303/se–2

Butt, A. K., Zubair, R., & Rathore, F. A. (2022). The role of augmentative and alternative communication in speech and language therapy: a mini review. *Journal of Pakistan Medical Association*, *72*(3), 581–584. https://doi.org/10.47391/jpma.22–023

Carroll, R. A., & Klatt, K. P. (2008). Using stimulus–stimulus pairing and direct reinforcement to teach vocal verbal behavior to young children with autism. *The Analysis of Verbal Behavior*, *24*(1), 135–146. https://doi.org/10.1007/bf03393062

Charlop-Christy, M. H., LeBlanc, L. A., & Carpenter, M. H. (1999). Naturalistic teaching strategies (NATS) to teach speech to children with autism: Historical perspective, development, and current practice. *The California School Psychologist*, *4*(1), 30-46. https://doi.org/10.1007/bf03340868

Chen, F., Wang, L., Peng, G., Yan, N., & Pan, X. (2019). Development and evaluation of a 3–D virtual pronunciation tutor for children with autism spectrum disorders. *PLOS ONE*, *14*(1), 1–22. https://doi.org/10.1371/journal.pone.0210858

Chenausky, K., Brignell, A., Morgan, A., & Tager–Flusberg, H. (2019). Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism & Developmental Language Impairments*, *4*, 1–12. https://doi.org/10.1177/2396941519856333

Chenausky, K., Norton, A., Tager–Flusberg, H., & Schlaug, G. (2016). Auditory–Motor Mapping Training: Comparing the effects of a novel speech treatment to a control treatment for minimally verbal children with autism. *PLOS ONE*, *11(11)*, 1–22. https://doi.org/10.1371/journal.pone.0164930

Chenausky, K., Norton, A., Tager–Flusberg, H., & Schlaug, G. (2018). Behavioral predictors of improved speech output in minimally verbal children with autism. *Autism Research*, *11*(10), 1356–1365. https://doi.org/10.1002/aur.2006

Chenausky, K. V., Norton, A. C., & Schlaug, G. (2017). Auditory–Motor Mapping Training in a more verbal child with autism. *Frontiers in Human Neuroscience*, *11*. 1–9. https://doi.org/10.3389/fnhum.2017.00426

Chericoni, N., de Brito Wanderley, D., Costanzo, V., Diniz–Gonçalves, A., Leitgel Gille, M., Parlato, E., Cohen, D., Apicella, F., Calderoni, S., & Muratori, F. (2016). Pre–linguistic vocal trajectories at 6–18 months of age as early markers of autism. *Frontiers in Psychology*, *7*, 1–11. https://doi.org/10.3389/fpsyg.2016.01595

Choi, J., Greer, R. D., & Keohane, D.–D. (2015). The effects of an auditory match–to–sample procedure on listener literacy and echoic responses. *Behavioral Development Bulletin*, *20*(2), 186–206. https://doi.org/10.1037/h0101313

Chumpelik, D. (1984). The PROMPT system of therapy: Theoretical framework and applications for developmental apraxia of speech. *Seminars in Speech and Language, 5*(2), 139–156. https://doi.org/10.1055/s-0028-1085172

Cividini–Motta, C., Scharrer, N., & Ahearn, W. H. (2017). An assessment of three procedures to teach echoic responding. *The Analysis of Verbal Behavior*, *33*(1), 41–63. https://doi.org/10.1007/s40616–016–0069–z

Cleland, J., Gibbon, F. E., Peppé, S. J. E., O'Hare, A., & Rutherford, M. (2010). Phonetic and phonological errors in children with high functioning autism and Asperger syndrome. *International Journal of Speech–Language Pathology*, *12*(1), 69–76. https://doi.org/10.3109/17549500903469980

Coon, J. T., & Miguel, C. F. (2012). The role of increased exposure to transfer-of-stimulus control procedures on the acquisition of intraverbal behavior. *Journal of Applied Behavior Analysis*, *45*(4), 657–666. https://doi.org/10.1901/jaba.2012.45-657

Council on Academic Accreditation in Audiology and Speech-Language Pathology. (2023). Standards for accreditation of graduate education programs in audiology and speech-language pathology. https://caa.asha.org/siteassets/files/accreditation-standards-for-graduate-programs.pdf

Dale, P. S., & Hayden, D. A. (2013). Treating speech subsystems in childhood apraxia of speech with tactual input: The PROMPT approach. *American Journal of Speech–Language Pathology*, *22*(4), 644–661. https://doi.org/10.1044/1058–0360(2013/12–0055)

Dawson, E. J. (2010). Current assessment and treatment practices for children with autism and suspected childhood apraxia of speech: A survey of speech–language pathologists (Publication No. 1479529) [Master's Thesis, Portland State University]. ProQuest Dissertations Publishing.

DeSouza, A. A., Akers, J. S., & Fisher, W. W. (2017). Empirical application of skinner's verbal behavior to interventions for children with autism: A Review. *The Analysis of Verbal Behavior*, *33*(2), 229–259. https://doi.org/10.1007/s40616–017–0093–7

Drash, P. W., High, R. L., & Tudor, R. M. (1999). Using mand training to establish an echoic repertoire in young children with autism. *The Analysis of Verbal Behavior*, *16*, 29–44. https://doi.org/10.1007/bf03392945

Du, L., Speckman, J., Medina, M., & Cole–Hatchard, M. (2017). The effects of an auditory matching iPad app on three preschoolers' echoic and listener responses. *Behavior Analysis in Practice*, *10*(2), 118–130. https://doi.org/10.1007/s40617–017–0174–z

Dyer, K., Santarcangelo, S., & Luce, S. C. (1987). Developmental influences in teaching language forms to individuals with developmental disabilities. *Journal of Speech and Hearing Disorders*, *52*(4), 335–347. https://doi.org/10.1044/jshd.5204.335

Dziuk, M. A., Larson, J. G., Apostu, A., Mahone, E. M., Denckla, M. B., & Mostofsky, S. H. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine & Child Neurology*, *49*(10), 734–739. https://doi.org/10.1111/j.1469–8749.2007.00734.x

Eikeseth, S., Smith, D., & Klintwall, L. (2014). Discrete trial teaching and discrimination training. In J. Tarbox, D.R. Dixon, P. Sturmey, & J. L. Matson (Eds.), *Handbook of early intervention for autism spectrum disorders: Research, practice, and policy* (pp. 229–253). Springer. https://doi.org/10.1007/978–1–4939–0401–3_12

Esch, B., Carr, J., & Michael, J. (2005). Evaluating stimulus–stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with autism. *The Analysis of Verbal Behavior*, *21*(1), 43–58. https://doi.org/10.1007/BF03393009

Fisher, W. W., Piazza, C. C., Bowman, L. G., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment to enhance reinforcer identification. *American Journal on Mental Retardation, 101*(1), 15–25.

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, *67*(3), 308–319. https://doi.org/10.1037/0022-006x.67.3.308

Freed, D.B., Marshall, R.C., & Frazier, K.E. (1997). Long-term effectiveness of PROMPT treatment in a severely apractic-aphasic speaker. *Aphasiology*, *11*(4/5). 365-372. https://doi.org/10.1080/02687039708248477

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Goldsmith, H. H. (2008). Infant and toddler oral– and manual–motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(1), 43–50. https://doi.org/10.1111/j.1469–7610.2007.01820.x

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183. https://doi.org/10.1017/S0305000900013167

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467–9280.2008.02117.x

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *44*(3), 475–498. https://doi.org/10.1901/jaba.2011.44–475

Hayden, D. (2004). PROMPT: A tactually grounded treatment approach to speech production disorders. In I.J. Stockman (Ed.), *Movement and Action in Learning and Development: Clinical Implications for Pervasive Developmental Disorders* (pp. 255–297). Elsevier Academic Press.

Hayden D. (2006). The PROMPT model: Use and application for children with mixed phonological–motor impairment. *Advances in Speech Language Pathology*, *8*(3), 265–281. https://doi.org/10.1080/14417040600861094

Hayden, D., Namasivayam, A. K., & Ward, R. (2015). The assessment of fidelity in a motor speech–treatment approach. *Speech, Language and Hearing*, *18*(1), 30–38. https://doi.org/10.1179/2050572814Y.0000000046

Hayden, D., Namasivayam, A. K., Ward, R., Clark, A., & Eigen, J. (2021). The PROMPT approach. In A.L. William, S. McLeod, & R.J. McCauley (Eds.), *Interventions for Speech Sound Disorders in Children* (2nd ed., pp. 477–504). Brookes Publishing.

Hegde, M. N., & Pena-Brooks, A. (2008). *Treatment protocols for articulation and phonological disorders*. Plural Publishing.

Henrickson, M. L., Rapp, J. T., & Ashbeck, H. A. (2015). Teaching with massed versus interspersed trials: Effects on acquisition, maintenance, and problem behavior: Massed versus interspersed trials. *Behavioral Interventions*, *30*(1), 36–50. https://doi.org/10.1002/bin.1396

Herreras Mercado, R., Simpson, K., & Bellom-Rohrbacher, K. H. (2019). Effect of prompts for restructuring oral muscular phonetic targets (PROMPT) on compensatory articulation in children with cleft palate/lip. *Global Pediatric Health*, *6*, 1–12. https://doi.org/10.1177/2333794X19851417

Hoff, E. (2001). *Language development* (2nd ed.). Wadsworth/Thomson Learning.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*(1), 185–241. https://doi.org/10.1901/jeab.1996.65-185

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *45*(2), 212–229. https://doi.org/10.1111/j.1469–7610.2004.00215.x

Jennett, H. K., Harris, S. L., & Delmolino, L. (2008). Discrete trial instruction vs. mand training for teaching children with autism to make requests. *The Analysis of Verbal Behavior*, *24*(1), 69–85. https://doi.org/10.1007/bf03393058

Kaufman, N. (1995a). *Psychometric properties and normative tables for the Kaufman Speech Praxis Test for Children*. Wayne State University Press.

Kaufman, N. (1995b). *Administering and scoring the Kaufman Speech Praxis Test for Children: Clinician's testing guide*. Wayne State University Press.

Kay, J. C., Kisamore, A. N., Vladescu, J. C., Sidener, T. M., Reeve, K. F., Taylor-Santa, C., & Pantano, N. A. (2020). Effects of exposure to prompts on the acquisition of intraverbals in children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, *53*(1), 493–507. https://doi.org/10.1002/jaba.606

Kjelgaard, M. M., & Tager–Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16*(2–3), 287–308. https://doi.org/10.1080/01690960042000058

Kodak, T., Fuchtman, R., & Paden, A. (2012). A comparison of intraverbal training procedures for children with autism. *Journal of Applied Behavior Analysis*, *45*(1), 155–160. https://doi.org/10.1901/jaba.2012.45-155

Koegel, L. K., Bryan, K. M., Su, P. L., Vaidya, M., & Camarata, S. (2020). Definitions of nonverbal and minimally verbal in research for autism: A systematic review of the literature. *Journal of Autism and Developmental Disorders*, *50*(8), 2957–2972. https://doi.org/10.1007/s10803–020–04402–w

Koegel, R. L., O'Dell, M., & Dunlap, G. (1988). Producing speech use in nonverbal autistic children by reinforcing attempts. *Journal of Autism and Developmental Disorders*, *18*(4), 525–538. https://doi.org/10.1007/BF02211871

Koenig, M., & Gerenser, J. (2006). SLP-ABA: Collaborating to support individuals with communication impairments. *The Journal of Speech and Language Pathology – Applied Behavior Analysis*, *1*(1), 2–10. https://doi.org/10.1037/h0100180

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, *100*(4), 2425–2438. https://doi.org/10.1121/1.417951

LaFrance, D. L., Weiss, M. J., Kazemi, E., Gerenser, J., & Dobres, J. (2019). Multidisciplinary teaming: Enhancing collaboration through increased understanding. *Behavior Analysis in Practice*, *12*(3), 709–726. https://doi.org/10.1007/s40617-019-00331-y

Leaf, J. B., Ferguson, J. L., & Cihon, J. H. (2022). Discrete trial teaching: Toward a progressive model. In J. B. Leaf, J. H. Cihon, J. L. Ferguson, & M. J. Weiss (Eds.), *Handbook of Applied Behavior Analysis Interventions for Autism* (pp. 171–184). Springer International Publishing. https://doi.org/10.1007/978–3–030–96478–8_10

Leaf, J. B., Leaf, R., McEachin, J., Taubman, M., Ala'i-Rosales, S., Ross, R. K., Smith, T., & Weiss, M. J. (2016). Applied behavior analysis is a science and, therefore, progressive. *Journal of Autism and Developmental Disorders*, *46*(2), 720–731. https://doi.org/10.1007/s10803-015-2591-6

Lee, C.C., Jhang, Y., Relyea, G., Chen, L., & Oller, D. K. (2018). Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all–day recordings. *Infant Behavior & Development*, *50*, 140–153. https://doi.org/10.1016/j.infbeh.2017.12.002

Lovaas, O. I., Berberich, J. P., Perloff, B. F., & Schaeffer, B. (1966). Acquisition of imitative speech by schizophrenic children. *Science, 151*(3711), 705–707.

Maestro, S., Muratori, F., Barbieri, F., Casella, C., Cattaneo, V., Cavallaro, M. C., Cesari, A., Milone, A., Rizzo, L., Viglione, V., Stern, D. D., & Palacio-Espasa, F. (2001). Early behavioral development in autistic children: The first 2 years of life through home movies. *Psychopathology*, *34*(3), 147–152. https://doi.org/10.1159/000049298

Maestro, S., Muratori, F., Cavallaro, M. C., Pecini, C., Cesari, A., Paziente, A., Stern, D., Golse, B., & Palacio-Espasa, F. (2005). How young children treat objects and people: An empirical study of the first year of life in autism. *Child Psychiatry and Human Development*, *35*(4), 383–396. https://doi.org/10.1007/s10578-005-2695-x

Maestro, S., Muratori, F., Cavallaro, M. C., Pei, F., Stern, D., Golse, B., & Palacio–Espasa, F. (2002). Attentional skills during the first 6 months of age in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(10), 1239–1245. https://doi.org/10.1097/00004703–200302000–00021

Maenner, M. J., Warren, Z., Williams, A.R., Amaokohene, E., Bakian, A.V., Bilder, D.A., Durkin, M.S., Fitzgerald, R.T., Furnier, S.M., Hughes, M.M., Ladd–Acosta, C.M., McArthur, D., Pas, E.T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall–Lande, J.,…Shaw, K.A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020. *Morbidity and Mortality Weekly Report Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Mars, A. E., Mauk, J. E., & Dowrick, P. W. (1998). Symptoms of pervasive developmental disorders as observed in prediagnostic home videos of infants and toddlers. *The Journal of pediatrics*, *132*(3), 500–504. https://doi.org/10.1016/s0022–3476(98)70027–7

Mayo, J., Chlebowski, C., Fein, D. A., & Eigsti, I.–M. (2013). Age of first words predicts cognitive ability and adaptive skills in children with ASD. *Journal of Autism and Developmental Disorders*, *43*(2), 253–264. https://doi.org/10.1007/s10803–012–1558–0

Mostofsky, S. H., Dubey, P., Jerath, V. K., Jansiewicz, E. M., Goldberg, M. C., & Denckla, M. B. (2006). Developmental dyspraxia is not limited to imitation in children with autism spectrum disorders. *Journal of the International Neuropsychological Society*, *12*(3), 314–326. https://doi.org/10.1017/s1355617706060437

Namasivayam, A. K., Coleman, D., O'Dwyer, A., & van Lieshout, P. (2020). Speech sound disorders in children: An articulatory phonology perspective. *Frontiers in Psychology*, *10*, 1-20. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02998

Namasivayam, A. K., Pukonen, M., Goshulak, D., Yu, V. Y., Kadis, D. S., Kroll, R., Pang, E. W., & De Nil, L. F. (2013). Relationship between speech motor control and speech intelligibility in children with speech sound disorders. *Journal of Communication Disorders*, *46*(3), 264–280. https://doi.org/10.1016/j.jcomdis.2013.02.003

Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics & Phonetics*, *20*(5), 351–369. https://doi.org/10.1080/02699200500211451

Norrelgen, F., Fernell, E., Eriksson, M., Hedvall, Å., Persson, C., Sjölin, M., Gillberg, C., & Kjellmer, L. (2015). Children with autism spectrum disorders who do not develop phrase speech in the preschool years. *Autism*, *19*(8), 934–943. https://doi.org/10.1177/1362361314556782

Oller, D. (1980). The emergence of the sounds of speech in infancy. In G.H. Yeni–Komshian, J.F. Kavanagh, & C.A. Ferguston (Eds.), *Child phonology: Volume1, production* (Vol. 1, pp. 73–92). Academic Press.

Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, *32*(4), 223–245. https://doi.org/10.1016/s0021–9924(99)00013–1

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, *14*(2), 239–251. https://doi.org/10.1017/S0954579402002031

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders*, *44*(10), 2413–2428. https://doi.org/10.1007/s10803–014–2047–4

Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: A two year follow–up. *Autism Research*, *1*(2), 97–107. https://doi.org/10.1002/aur.12

Pecukonis, M., Plesa Skwerer, D., Eggleston, B., Meyer, S., & Tager–Flusberg, H. (2019). Concurrent social communication predictors of expressive language in minimally verbal children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *49*(9), 3767–3785. https://doi.org/10.1007/s10803–019–04089–8

Plumb, A. M., & Wetherby, A. M. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, *56*(2), 721–734. https://doi.org/10.1044/1092–4388

Preston, D., & Carter, M. (2009). A review of the efficacy of the picture exchange communication system intervention. *Journal of Autism and Developmental Disorders*, *39*(10), 1471–1486. https://doi.org/10.1007/s10803-009-0763-y

Ramsdell–Hudock, H. L., Warlaumont, A. S., Foss, L. E., & Perry, C. (2019). Classification of infant vocalizations by untrained listeners. *Journal of Speech, Language, and Hearing Research*, *62*(9), 3265–3275. https://doi.org/10.1044/2019_JSLHR–S–18–0494

Rapin, I., Dunn, M., Allen, D. A., Stevens, M., & Fein, D. (2009). Subtypes of language disorders in school–age children with autism. *Developmental Neuropsychology,34*(1), 66–84. https://doi.org/10.1080/87565640902769295

Roche, L., Zhang, D., Bartl–Pokorny, K. D., Pokorny, F. B., Schuller, B. W., Esposito, G., Bölte, S., Roeyers, H., Poustka, L., Gugatschka, M., Waddington, H., Vollmann, R., Einspieler, C., & Marschik, P. B. (2018). Early vocal development in autism spectrum disorder, Rett syndrome, and fragile x syndrome: Insights from studies using retrospective video analysis. *Advances in Neurodevelopmental Disorders*, *2*(1), 49–61. https://doi.org/10.1007/s41252–017–0051–3

Rogers, S. J., Hayden, D., Hepburn, S., Charlifue–Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver Model and PROMPT interventions. *Journal of Autism & Developmental Disorders*, *36*(8), 1007–1024. https://doi.org/10.1007/s10803–006–0142–x

Romski, M., & Sevcik, R. A. (2005). Augmentative communication and early intervention: Myths and realities. *Infants & Young Children*, *18*(3), 174–185. https://doi.org/10.1097/00001163–200507000–00002

Roncati, A. L., Souza, A. C., & Miguel, C. F. (2019). Exposure to a specific prompt topography predicts its relative efficiency when teaching intraverbal behavior to children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, *52*(3), 739–745. https://doi.org/10.1002/jaba.568

Rose, V., Trembath, D., Keen, D., & Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community–based early intervention programme. *Journal of Intellectual Disability Research*, *60*(5), 464–477. https://doi.org/10.1111/jir.12284

Roth, F. P., & Worthington, C. K. (2018). *Treatment resource manual for speech–language pathology* (5th ed.). Plural Publishing.

Saul, J., & Norbury, C. (2020). Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment? *Autism*, *24*(5), 1217–1231. https://doi.org/10.1177/1362361319898560

Seaver, J. L., & Bourret, J. C. (2014). An evaluation of response prompts for teaching behavior chains. *Journal of Applied Behavior Analysis*, *47*(4), 777–792. https://doi.org/10.1002/jaba.159

Shriberg, L. D. (2010). Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In R. Paul & P. Flipsen (Eds.), *Speech sound disorders in children* (pp. 1–33). Plural Publishing.

Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(4), 405–426. https://doi.org/10.1007/s10803–010–1117–5

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high–functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1097–1115. https://doi.org/10.1044/1092–4388

Shriberg, L.D., Strand, E.A., Jakielski, K.J., & Mabie, H.L. (2019) Estimates of the prevalence of speech and motor speech disorders in persons with complex neurodevelopmental disorders. *Clinical Linguistics & Phonetics, 33*(8), 707–736. https://doi.org/10.1080/02699206.2019.1595732

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*, *64*(1), 1–142. https://doi.org/10.1111/mono.1999.64.issue–1

Siller, M., & Sigman, M. (2008). Modeling longitudinal change in the language abilities of children with autism: Parent behaviors and child characteristics as predictors of change. *Developmental Psychology*, *44*(6), 1691–1704. https://doi.org/10.1037/a0013771

Skinner, B.F. (1957). *Verbal behavior*. Prentice-Hall.

Smith, V., Mirenda, P., & Zaidman–Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research*, *50*(1), 149–160. https://doi.org/10.1044/1092–4388

Speckman–Collins, J., Lee Park, H.–S., & Greer, R. D. (2007). Generalized selection–based auditory matching and the emergence of the listener component of naming. *Journal of Early and Intensive Behavior Intervention*, *4*(2), 412–429. https://doi.org/10.1037/h0100382

Stark, R. (1980). Stages of speech development in the first year of life. In G.H. Yeni–Komshian, J.F. Kavanagh, & C.A. Ferguston (Eds.), *Child phonology: Volume1, production* (Vol. 1, pp. 73–92). Academic Press.

Stone, W. L., & Yoder, P. J. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism*, *5*(4), 341-361. https://doi.org/10.1177/1362361301005004002

Su, P. L., Rogers, S. J., Estes, A., & Yoder, P. (2021). The role of early social motivation in explaining variability in functional language in toddlers with autism spectrum disorder. *Autism*, *25*(1), 244–257. https://doi.org/10.1177/1362361320953260

Sundberg, M. L. (2015, December 30). The most important verbal operant. *VB News*. https://marksundberg.com/wp-content/uploads/2019/01/T-Sundberg-2015.pdf

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities*. AVB Press.

Tager–Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., Rice, M., Stoel–Gammon, C., Wetherby, A., & Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *52*(3), 643–652. https://doi.org/10.1044/1092–4388(2009/08–0136)

Tager–Flusberg, H., & Kasari, C. (2013). Minimally verbal school–aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, *6*(6), 468–478. https://doi.org/10.1002/aur.1329

Tarbox, J., Madrid, W., Aguilar, B., Jacobo, W., & Schiff, A. (2009). Use of chaining to increase complexity of echoics in children with autism. *Journal of Applied Behavior Analysis*, *42*(4), 901–906. https://doi.org/10.1901/jaba.2009.42–901

Taylor, B. A., & Fisher, J. (2010). Three important things to consider when starting intervention for a child diagnosed with autism. *Behavior Analysis in Practice*, *3*(2), 52–53. https://doi.org/10.1007/BF03391765

Terband, H., Maassen, B., & Maas, E. (2019). A psycholinguistic framework for diagnosis and treatment planning of developmental speech disorders. *Folia Phoniatrica et Logopaedica*, *71*(5–6), 216–227. https://doi.org/10.1159/000499426

The PROMPT Institute. (2009). Introduction to technique: A manual.

Thurm, A., Lord, C., Lee, L.–C., & Newschaffer, C. (2007). Predictors of language acquisition in preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(9), 1721–1734. https://doi.org/10.1007/s10803–006–0300–1

Tierney, C., Mayes, S., Lohs, S., Black, A., Gisin, E., & Veglia, M. (2015). How valid is the checklist for autism spectrum disorder when a child has apraxia of speech? *Journal of Developmental and Behavioral Pediatrics*, *36*(8), 569–574. https://doi.org/10.1097/DBP.0000000000000189

Velleman, S. L., Andrianopoulos, M. V., Boucher, M., Perkins, J., Averback, K., Currier, A. R., & Van Emmerik, R. (2009). Motor speech disorders in children with autism. In R. Paul & P. Flipsen (Eds.), *Speech sound disorders in children* (pp. 141–180). Plural Publishing.

Venter, A., Lord, C., & Schopler, E. (1992). A follow–up study of high–functioning autistic children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *33*(3), 489–507. https://doi.org/10.1111/j.1469–7610.1992.tb00887.x

Verschuur, R., Didden, R., Lang, R., Sigafoos, J., & Huskens, B. (2013). Pivotal response treatment for children with autism spectrum disorders: A systematic review. *Review Journal of Autism and Developmental Disorders*, *1*. https://doi.org/10.1007/s40489-013-0008-z

Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Truemper, K., Rusiewicz, H. L., & Moore, C. A. (2014). Data–driven subclassification of speech sound disorders in preschool children. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2033–2050. https://doi.org/10.1044/2014_JSLHR–S–12–0193

Wan, C. Y., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., Norton, A., & Schlaug, G. (2011). Auditory–motor mapping training as an intervention to facilitate speech output in non–verbal children with autism: A proof of concept study. *PLoS ONE*, *6*(9), e25505. https://doi.org/10.1371/journal.pone.0025505

Ward, R., Strauss, G., & Leitão, S. (2013). Kinematic changes in jaw and lip control of children with cerebral palsy following participation in a motor-speech (PROMPT) intervention. *International Journal of Speech-Language Pathology*, *15*(2), 136–155. https://doi.org/10.3109/17549507.2012.713393

Ward, R., Leitão, S., & Strauss, G. (2014). An evaluation of the effectiveness of PROMPT therapy in improving speech production accuracy in six children with cerebral palsy. *International Journal of Speech-Language Pathology*, *16*(4), 355–371. https://doi.org/10.3109/17549507.2013.876662

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

Weaver, A. D., McKevitt, B. C., & Farris, A. M. (2017). Using multiple–stimulus without replacement preference assessments to increase student engagement and performance. *Beyond Behavior*, *26*(1), 5–10. https://doi.org/10.1177/1074295617694410

Werner, E., & Dawson, G. (2005). Validation of the phenomenon of autistic regression using home videotapes. *Archives of General Psychiatry*, *62*(8), 889–895. https://doi.org/10.1001/archpsyc.62.8.889

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one Year of age: A retrospective study based on home videotapes. *Journal of Autism and Developmental Disabilities, 30*(2), 157-162.

Wetherby, A., Chambers, N., Morgan, L., & Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders*, *37*, 960–975. https://doi.org/10.1007/s10803–006–0237–4

Xu, S., Lee, G. T., Feng, H., Niu, R., Gilic, L., & Shao, W. (2020). Comparing the effects of echoic prompts and echoic plus picture prompts on establishing intraverbal behavior for children with autism. *Education and Training in Autism and Developmental Disabilities*, *55*(1), 89-100. https://www.jstor.org/stable/26898716

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value–added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *45*(5), 1254–1270. https://doi.org/10.1007/s10803–014–2286–4

Yu, V. Y., Kadis, D. S., Oh, A., Goshulak, D., Namasivayam, A., Pukonen, M., Kroll, R., De Nil, L. F., & Pang, E. W. (2014). Changes in voice onset time and motor speech skills in children following motor speech therapy: Evidence from /pa/ productions. *Clinical Linguistics & Phonetics*, *28*(6), 396–412. https://doi.org/10.3109/02699206.2013.874040

Yu, V.Y., Kadis, D. S., Goshulak, D., Namasivayam, A. K., Pukonen, M., Kroll, R. M., De Nil, A.F., & Pang, E. W. (2018). Impact of motor speech intervention on neural activity in children with speech sound disorders: Use of magnetoencephalography. *Journal of Behavioral and Brain Science*, *8*(7), 415-429. https://doi.org/10.4236/jbbs.2018.87026

Zablotsky, B., Black, L.I., Maenner, M.J., Schieve, L.A., Danielson, M.L., Bitsko, R.H., Blumberg, S.J., Kogan, M.D., & Boyle, C.A. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics*, *144*(4), 1–11. https://doi.org/10.1542/peds.2019–0811

Zappella, M., Einspieler, C., Bartl-Pokorny, K. D., Krieber, M., Coleman, M., Bölte, S., & Marschik, P. B. (2015). What do home videos tell us about early motor and socio-communicative behaviours in children with autistic features during the second year of life – an exploratory study. *Early Human Development*, *91*(10), 569–575. https://doi.org/10.1016/j.earlhumdev.2015.07.006

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696