Cross-modal generalization of receptive and expressive vocabulary in children with autism spectrum disorder

Pumpki L Su, George Castle and Stephen Camarata
Vanderbilt University Medical Center, USA

Abstract

Background and aims: Word learning is an area that poses a particular challenge to children with autism spectrum disorder. A unique challenge for this population is generalization of learned skills across new learning contexts. In clinical settings, a common assumption in teaching vocabulary for children with autism spectrum disorder is that learning in one modality will generalize incidentally to untreated modalities, but very few studies have evaluated the validity of this assumption. The purpose of this study was to investigate receptive and expressive word acquisition and cross-modal generalization in children with autism spectrum disorder.

Methods: A single-case parallel treatments design was used to compare word learning and cross-modal generalization in children with autism spectrum disorder. Ten children with autism spectrum disorder were taught unfamiliar vocabulary words in a combined storybook and play intervention. For each child, half of the target words were trained expressively and the other half were trained receptively by random assignment. No direct cross-modal instruction was delivered. A series of probe sessions were completed to assess participants’ within-modal learning and cross-modal generalization of vocabulary learning.

Results: All children learned target words in both receptive and expressive conditions, as evidenced by an average of 80% accuracy across three trials at the end of each intervention. Overall, cross-modal generalization was higher for the expressive-to-receptive direction than for the receptive-to-expressive direction. Nine out of ten children demonstrated successful cross-modal generalization on the expressive-to-receptive probes and only three children demonstrated successful cross-modal generalization on the receptive-to-expressive probes. Large variability was observed with regard to number of intervention sessions needed to reach mastery criterion and there were individual patterns of word learning.

Conclusion: Contrary to the assumption that vocabulary learning will be “automatically” generalized across modalities, results from this study indicate that cross-modal generalization at the word level is not automatic nor consistent in children with autism spectrum disorder, particularly in the receptive-to-expressive direction.

Implications: The finding that more children demonstrated expressive-to-receptive generalization than the opposite direction suggests that targeting expressive vocabulary first with the goal of incidentally increasing receptive vocabulary may be more efficient than starting with the receptive modality. Additionally, the finding that not all children demonstrated successful expressive-to-receptive cross-modal generalization indicates that teaching vocabulary in the expressive modality exclusively does not guarantee receptive understanding in this population. We recommend that practitioners periodically monitor children’s vocabulary learning in both modalities or set an explicit generalization goal to ensure complete learning of trained words.

Keywords
Autism spectrum disorders, communication and language, lexical development, intervention/therapy, single case design

Corresponding author:
Stephen Camarata, Department of Hearing and Speech Science, Vanderbilt University Medical Center, Nashville, USA.
Email: stephen.camarata@vumc.org

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Language deficits are a core feature of autism spectrum disorder (ASD), and children with ASD demonstrate these language deficits early on in life (Ellis Weismer, Lord, & Esler, 2010; Paul, Chawarska, Cicchetti, & Volkmar, 2008). There is an extensive literature indicating that vocabulary learning is an area that poses a particular challenge to this population (Arunachalam & Luyster, 2016; Charmann, Drew, Baird, & Baird, 2003; Herlihy, Knoch, Vibert, & Fein, 2015; Luyster, Lopez, & Lord, 2007; Luyster & Lord, 2009). Indeed, delayed onset of first words is often the first recognized symptom in ASD (Camarata, 2014; De Giacomo & Fombonne, 1998). Receptive language delay can be observed as early as 12 months in ASD and expressive vocabulary delay at 18 months (Mitchell et al., 2006). Children with ASD also experience significant delays in both expressive and receptive modalities of word learning (Hudry et al., 2010). Reduced word learning proficiency is witnessed in a substantial proportion of the ASD population (Hus, Pickles, Cook, Risi, & Lord, 2007). Given that vocabulary acquisition is strongly linked to later language skills in typically developing (TD) children and children with ASD (Howlin, Mawhood, & Rutter, 2000), additional research to understand word learning deficits and effective strategies to facilitate word learning in this population is warranted.

Children with ASD display a unique vocabulary learning profile in two aspects. First, they present with a larger weakness in receptive language than expected relative to their expressive language (Kover, McDuffie, Hagerman, & Abbeduto, 2013; Luyster, Kadlec, Carter, & Tager-Flusberg, 2008; Mitchell et al., 2006). In typical development, receptive vocabulary development usually precedes expressive vocabulary development in absolute terms and is strongly correlated with expressive language (Bornstein & Hendricks, 2012). The relative weakness in receptive vocabulary (as compared to expressive vocabulary) in ASD has been reported across age groups (Ellis Weismer et al., 2010; Kover et al., 2013; Pickles, Anderson, & Lord, 2014) and divergent cognitive profiles (Kjelgaard & Tager-Flusberg, 2001; Woynaroski, Yoder, & Watson, 2016). The atypical expressive-receptive vocabulary profile is not only evident cross-sectionally but can also be observed longitudinally: several longitudinal studies have confirmed that the longitudinal association between receptive and expressive language is atypical in infants and preschoolers with ASD (Hudry et al., 2014; Woynaroski et al., 2016). There has been additional evidence from a recent longitudinal study that this discrepant vocabulary profile gradually changes over development across the preschool period, suggesting that it may be a particular relevant marker for young children with ASD (Davidson & Ellis Weismer, 2017).

Second, another unique challenge for children with ASD is generalization of learned skills across materials, settings, or communication partners (Eldevik, Kazemi, & Elsky, 2016). For example, children and adolescents with ASD generalized learned word learning strategies less consistently to a new learning context compared to TD participants (de Marchena, Eigsti, & Yerys, 2015) and failed to generalize learned language skills across social communication partners (Pellecchia & Hineline, 2007). Difficulties with generalization have long been reported in the literature in ASD and have persisted despite advances in intervention (Lovaas, Koegel, Simmons, & Long, 1973; Tek & Naigles, 2017).

Cross-modal generalization

Cross-modal generalization is one specific type of generalization that children with ASD may have difficulty with even relative to other populations with language disabilities (e.g., intellectual disabilities). Cross-modal generalization is defined as the transfer of learned vocabulary from one modality (e.g. expressive vocabulary) to another modality (e.g. receptive vocabulary) without direct instruction in the latter “cross” modality (Davis, Camarata, & Camarata, 2016a). TD children generalize cross-modally at the word level for both receptive and expressive modalities without explicit training and often do so quite rapidly (Dollaghan, 1985; Fernandes, 2008). This ability to generalize learned skills to an untrained modality was also termed as “cross-over effect” in early literature (Bucher & Keller, 1981; Guess, 1969). In Guess’s early study, cross-modal generalization was not found in either receptive-to-expressive or expressive-to-receptive directions in participants with developmental disabilities.

In contrast, more recent work has suggested that children with developmental disabilities and language impairments are capable of cross-modal generalization but are more likely to generalize from the expressive modality to receptive than vice versa (Davis, Lancaster, & Camarata, 2016b; Doyle, Wolery, Gast, Ault, & Wiley, 1990). In an intervention study that targeted expressive morphosyntax skills in children with specific language impairment (SLI), participants were found to display significant gains in receptive morphosyntax skills as incidental learning from the expressive treatment (Camarata, Nelson, Gillum, & Camarata, 2009). Specific to word learning, children with Down Syndrome and children with SLI similarly demonstrated some evidence of expressive-to-receptive generalization but consistent difficulty with receptive-to-expressive generalization compared to mental age-matched TD peers (Bird, Chapman, & Schwartz, 2004; Davis et al., 2016a; Gray, 2003). Regarding
receptive-to-expressive generalization, only one study to date (Bucher & Keller, 1981) has provided evidence that receptive-to-expressive generalization is possible in children with language impairments without cognitive deficits but is influenced by factors such as complexity of target vocabulary, familiarity of stimuli, and the extent to which participants learned the vocabulary. Participants demonstrated enhanced receptive-to-expressive generalization with simpler and shorter vocabulary, more familiar target word pictures, and longer training in the receptive modality. These findings collectively suggest that cross-modal generalization cannot be assumed in children with different language impairment etiologies, especially in the receptive-to-expressive direction.

**Gap in current literature on teaching vocabulary to children with ASD**

In clinical settings, professionals continue to debate about when and in which sequence we should teach vocabulary when working with children with ASD. Based on a common vocabulary acquisition sequence in TD children, some researchers have recommended a receptive-before-expressive intervention sequence (Lovaas, 2003; Taylor & McDonough, 1996). However, a review by Petursdottir and Carr (2011) suggests that empirical support for this recommendation is limited. An alternative approach is to target vocabulary in the expressive modality before the receptive modality or only in the expressive modality. In fact, numerous intervention packages for children with ASD were designed to focus on language production and a substantial portion of intervention studies in this population examined the effects of treatment on expressive language as primary outcomes (i.e. verbalization, word production, initiation of communication, etc.) and receptive language as secondary outcomes (see Goldstein, 2002 for a review and Hampton & Kaiser, 2016 for a recent meta-analysis).

However, the assumption that teaching vocabulary in the expressive modality will improve receptive vocabulary in the population of ASD has only been tested in one study (Wynn & Smith, 2003). In this study, though targeting vocabulary in the expressive-only condition resulted in cross-modal generalization more often than targeting vocabulary in the receptive-only condition, not all children consistently demonstrated expressive-to-receptive generalization. Additionally, the vocabulary targets used in Wynn and Smith may have been quite simple for participants to learn. The number of training sessions in each condition ranged from one to five sessions for each word pair, which may have been too few to demonstrate conclusive within-subject comparisons. The reviewed literature shows that evidence for teaching expressive and receptive vocabulary in children with ASD is extremely limited. On one hand, past work on cross-modal generalization in children with developmental disabilities and language impairments suggests that children with ASD will likely display a similar expressive-to-receptive generalization pattern. On the other hand, given the distinctive expressive-receptive vocabulary profile and particular difficulty in generalization in children with ASD, there is some evidence to support that children with ASD may experience difficulty generalizing in both directions (Wynn & Smith, 2003). Taken together, it remains unclear whether children with ASD are capable of incidentally generalizing vocabulary learned in one modality to another.

**The current study**

Vocabulary is a foundation for many aspects of language development and is an important predictor of later academic success in children with typical development and with language impairments (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003). Studying cross-modal generalization has the potential to help understand specific word learning deficits in ASD and increase cost efficiency in language intervention. A compelling practical rationale is that if children with ASD are capable of making gains in receptive vocabulary as an incidental effect of expressive-only language interventions, then expressive-only language interventions may be an efficient and sufficient program for teaching vocabulary. On the other hand, it is important to directly test whether teaching vocabulary in either the expressive or receptive modality will increase learning efficiency via incidental, untrained cross-modal generalization. The purpose of this study was to investigate receptive and expressive word acquisition and cross-modal generalization in children with ASD. Specifically, we asked the following research questions:

1. Do children with ASD generalize cross-modally from expressively-only trained targets to untrained receptive target identification following a combined storybook and play expressive vocabulary intervention?
2. Do children with ASD generalize cross-modally from receptively-only trained targets to untrained expressive target identification following a combined storybook and play receptive vocabulary intervention?
3. Does expressive-only vocabulary instruction lead to greater generalization as measured by a higher percentage of correct responses in cross-modal probes than receptive-only vocabulary instruction?
4. Does expressive-only vocabulary instruction lead to faster learning as measured by fewer sessions to
reach mastery criterion in within-modal probes than receptive-only vocabulary instruction?

Methods

Data were collected between June 2009 and March 2011. The University Institutional Review Board reviewed and approved all study procedures. A parent or legal guardian provided informed consent for all participants. Assent was not obtained because participants’ low language levels precluded a plausible case that assent narratives could be adequately comprehended. Rather, specific behavioral guidelines were adopted in order to ensure that a participant could reasonably demonstrate their wish to terminate the study at any phase (and in any session).

Participants

Participants included nine boys with ASD and one girl with ASD, ages 3.0–7.4 years with a mean age of 4.6 years (SD = 1.5 years). All participants were recruited through local public schools and the university clinic. Participants were identified through parent and clinician reports as qualifying for language intervention services and validated by subsequent standardized testing. Inclusionary criteria were as follows: (a) age of 36–90 months, (b) an established diagnosis of ASD through medical records, (c) ability to imitate an adult’s vocal behavior, and (d) availability to commit to enrollment in intervention for six to eight months.

We confirmed all participants’ diagnosis of ASD using the Autism Diagnostic Observation Schedule–Generic (Lord et al., 2000). Participants’ cognitive abilities were measured using the Revised Leiter International Performance Scale (Leiter-R; Roid & Miller, 2011). Participants’ IQ ranged from 67 to 102, with a mean of 88 (SD = 12). An omnibus language assessment, the Preschool Language Scale-3 (Zimmerman, Steiner, & Pond, 1992), was used to measure general language ability. Additionally, participants’ receptive language was measured using the Test of Auditory Language Comprehension-3 (Carrow-Woolfolk, 1999) and the Peabody Picture Vocabulary Test-4 (Dunn & Dunn, 2007). Expressive language was measured using the Expressive One-Word Picture Vocabulary Test-4 (Martin & Brownell, 2011). Demographic characteristics and standardized assessment results are summarized in Table 1.

Target vocabulary selection

A four-step process was employed to identify a set of 16 individualized target words that were absent from each child’s receptive and expressive vocabulary repertoire. First, a pool of vocabulary was selected for each child’s parent to indicate which words were absent from their child’s repertoire. The pool of words included low-incidence vocabulary words from the upper levels of the MacArthur-Bates Communicative Development Inventories (Fenson, Marchman, Thal, Dale, & Reznick, 2007), such as toucan, abacus, canoe, flashlight, tepee, or blimp. Second, the experimenter engaged each child in a word imitation task. Words for which the child could not produce an intelligible imitation were excluded so that a child’s phonological limitations would not confound expressive production or generalization. Third, the child was instructed to match one picture for each of the vocabulary words to an identical picture. The experimenter did not name the vocabulary during the matching task. Only the stimuli that the child correctly matched were included. This step ensured that the child had the necessary prerequisite matching skills for receptive probes so that errors in receptive probes would not be confounded by lack of ability to match or point to a target vocabulary. Fourth, the child was asked to identify each potential target word expressively and receptively. To test for expressive knowledge, the child was asked to name each picture. To test for receptive knowledge, the child was asked to identify each picture from a closed set of four pictures. Only vocabulary items that were not named correctly (0% across four expressive opportunities) and that were identified at or below chance level (< 25% across four receptive opportunities) were selected for intervention.

These procedures were important because we wanted to select vocabulary items that were absent from each child’s expressive and receptive repertoire and eliminate potential confounding effects of nonexperimental variables on demonstrating vocabulary learning such as cognitive ability to match pictures, motor ability to point to pictures, or oral-motor ability to express a target form intelligibly. After 16 vocabulary items were selected following these procedures, they were randomly assigned into four sets. Within each set, the words were randomly assigned to a modality so that two words within each set were taught expressively exclusively and two words were taught receptively exclusively.

Three sets were presented during intervention phases and a fourth set served as an untreated control set to monitor possible incidental learning of vocabulary and maturation effect. Control words were not taught in either modality but were included during each probe phase. Baseline procedures included at least three sessions of probes on all 16 words (targeted and control) to establish a stable data pattern before the intervention phase was implemented.
Experimental design

A single-case parallel treatments design across behavior sets (Gast & Ledford, 2018; Gast & Wolery, 1988) was used to compare word learning and cross-modal generalization of words learned within the receptive-only and expressive-only vocabulary intervention. This design is well suited for comparing interventions on nonreversible behaviors (Gast & Ledford, 2018). Conceptually, this design is equivalent to two concurrent multiple-probe designs implemented simultaneously with one variation of the intervention (Storytime Intervention + expressive prompts) implemented in one multiple-probe design whereas the second variation (Storytime Intervention + receptive prompts) in the other (Wolery, 2013). Experimental control was demonstrated through the staggered introduction of the independent variable to different sets of vocabulary with changes in learning behavior observed only after the introduction of the independent variable (Gast & Ledford, 2018).

One advantage of the parallel treatments design is that it does not require continuous baseline probe sessions for behaviors that have yet to be introduced to the intervention (Gast & Ledford, 2018; Jones & Schwartz, 2004). Therefore, it allows demonstration and comparison of two interventions for teaching nonreversible behaviors by time lagging the behavior sets (i.e. target word sets in this study) while controlling for maturation and history threats and avoiding instrumentation threats that are common in multiple baseline designs. This design also allowed expressive and receptive vocabulary to be targeted simultaneously yet independently so that word learning and generalization can be compared directly across two modalities. By “independent,” we are not arguing that vocabulary can be learned or taught completely in one modality in isolation. From an applied behavioral perspective on teaching communication and language, the behavior of the teacher and the behavior of the learner are functionally independent of one another (Sundberg & Michael, 2001; Sundberg & Partington, 1998). We are using the terms “isolation” or “independence” in modality to refer to the modality of the teaching prompts used in the intervention.

All children participated in receptive- and expressive-only vocabulary interventions and were taught vocabulary words that were absent in baseline expressive probes and at or below chance levels in baseline receptive probes. Vocabulary intervention modality (expressive or receptive) was randomly assigned to target words so that half of the target words were trained only expressively (expressive-only condition) and half of the words were trained only receptively (receptive-only condition) for each participant. No direct cross-modal training was delivered. A series of probe sessions were completed to assess baseline, within-modal learning, and cross-modal generalization of vocabulary learning. These probe sessions occurred during baseline assessment, during intervention, and post-intervention. Across both expressive- and receptive-only conditions, participants were trained to reach the same correction criterion in the targeted modality (80% correct responses across three sessions) before they were probed in the untrained modality. The primary dependent variable was the percentage of correct responses on expressive and receptive cross-modal probes. As an exploratory analysis, our secondary dependent variable was the number of sessions needed to learn vocabulary to criterion (defined later in this section). A difference in the effectiveness of the expressive-only condition versus receptive-only condition

| Participant | Age (months) | Sex | Leiter-R | PLS-3 | TACL-3 | PPVT-4 | EOWPVT-4 |
|-------------|--------------|-----|----------|-------|--------|--------|----------|
| 1           | 39           | M   | 95       | 71    | 81     | 82     | 94       |
| 2           | 88           | M   | 80       | <50a  | 53     | 56     | 57       |
| 3           | 68           | M   | 67       | 50    | 55     | 53     | <55a     |
| 4           | 36           | M   | 74       | 69    | 68     | 67     | 65       |
| 5           | 76           | M   | 93       | 50    | 51     | 49     | 57       |
| 6           | 60           | F   | 85       | 52    | 61     | 66     | <55a     |
| 7           | 49           | M   | 102      | 63    | 79     | 85     | 77       |
| 8           | 45           | M   | 102      | 58    | 89     | 83     | 62       |
| 9           | 48           | M   | 82       | 56    | 66     | 53     | 55a      |
| 10          | 40           | M   | 100      | 85    | 102    | 103    | 108      |

EOWPVT-4: Expressive One-Word Picture Vocabulary Test—Fourth Edition (Martin & Brownell, 2011); Leiter–R: Revised Leiter International Performance Scale (Roid & Miller, 2011); PLS-3: Preschool Language Scales—Third Edition (Zimmerman et al., 1992); PPVT-4: Peabody Picture Vocabulary Test—Fourth Edition (Dunn & Dunn, 2007); TACL-3: Test of Auditory Language Comprehension—Third Edition (Carrow-Woolfolk, 1999). *Participants with scores below floor were entered as floor.
would be indicated if one modality of vocabulary instruction reliably produced greater cross-modal generalization or more vocabulary learning in a shorter amount of time than the contrasting modality.

**Vocabulary intervention**

Individual intervention sessions were conducted at the university clinic for each participant by licensed speech-language pathologists. For all participants, intervention started within one month of initial consent and assessments. Participants attended multiple 24-minute treatment sessions per week with a range of two to four sessions per week based on their ability to attend the sessions. The intervention length varied for each participant with a mean of 19 weeks, ranging from 14 weeks to 36 weeks. A combined storybook and play intervention using milieu teaching strategies (Storytime Intervention; Wolery, Ault, & Doyle, 1992a; Wolery et al., 1992b) was adopted. Each session was divided equally into an expressive-only condition and a receptive-only condition (12 minutes in each modality).

The clinician, intervention setting, play context, and reinforcers were identical across the expressive- and receptive-only conditions. The clinician introduced targeted vocabulary words in a storybook followed by structured play. During the storybook session, the child and the clinician viewed a picture book of at least 10 pages together while the clinician told a story about the pictures. Each target word was verbally presented by the clinician 10 times. During the play context, the clinician and the child interacted with a set of toys that matched the storybook theme. The clinician used commenting play strategy in which she used target words meaningfully for approximately 10 times after securing the child’s attention. The clinician also used other naturalistic play techniques, including environmental arrangement, shadow play, semantic expansion, and matched turn taking (Wolery et al., 1992a).

The expressive- and receptive-only interventions differed in the type of prompts provided by the clinician during intervention. For the expressive-only condition, the clinician prompted the child to name the vocabulary item when shown a picture that matched the word immediately after the clinician produced each vocabulary word. For the receptive-only condition, the clinician prompted the child to point to a picture that matched the word.

The clinician reinforced correct responses with verbal praises, smiles, or social routines such as high-fives. For incorrect responses, the clinician delivered the correct response (“Uh-oh, it’s a ____” for the expressive-only condition and “Uh-oh, here’s the ____” while pointing to the correct picture for the receptive-only condition) and readministered the prompt. If the child then identified or produced the word correctly, the clinician reinforced the correct response. If the child made no response after prompting, the clinician waited for a minimum of 2 seconds and then delivered the correct expressive or receptive exemplar.

**Probe procedures**

**Baseline probes.** At baseline, there were a minimum of three within-modal probe sessions for each participant to establish a stable data pattern before initiating the intervention phase. Each word was probed in the same modality that it was assigned to for intervention. Three within-modal probe trials of each word were included in each probe session. The criterion of moving from baseline phase to initial intervention phase was 0% of correct answers for expressive probes and less than 25% of correct answers (i.e. less than chance) for receptive probes across baseline probe sessions. During each within-modal probe session, the child was asked to point to pictures representing the targeted vocabulary words for words assigned to the receptive condition and to name the targeted vocabulary for words assigned to the expressive condition. The order of the probe trials was randomized for each participant. To maintain child motivation and engagement, familiar words that the child knew per parent report (e.g. “horse”) were interspersed with the target vocabulary probes. No reinforcement was delivered during probe sessions.

In addition to the within-modal probes, cross-modal probes (i.e. untrained modality) were also administered at baseline during one cross-modal probe session. For example, if the word “pelican” was assigned to the receptive-only condition for a child, it would be probed expressively during cross-modal probe sessions. All target words, including trained, untrained, and control sets of vocabulary were included in the cross-modal probe session. Similar to within-modal probe sessions, three trials of each word were included in each cross-modal probe session.

**Daily intervention probes.** After the completion of initial baseline probe sessions, the intervention began for the first set of four words. Two words were treated in the receptive-only condition, whereas the other two were treated in the expressive-only condition. The order in which receptive and expressive condition occurred was counterbalanced across clinic visits. Daily intervention probes occurred at the beginning of each clinic visit to track children’s within-modal word learning. Only words targeted in each phase were included in daily intervention probes. These words were only probed in the treated modality. The percentage of correct responses were collected, plotted after each session, and used to inform phase change. After the child
reached 80% correct responses on daily intervention probes across three consecutive sessions for all words in each set, the post-intervention probe sessions were initiated.

**Post-intervention probes.** Post-intervention probe sessions were identical to baseline probe sessions. Following intervention for each set of target vocabulary, three within-modal probe sessions and one cross-modal probe session took place.

### Measuring mastery and generalization

Participants’ responses to within-modal, cross-modal, and daily intervention probes were recorded during each probe or intervention session. The percentage of correct responses to probes was calculated and plotted for each word and each participant for visual analysis of data. The criterion for mastery of within-modal learning was defined as 80% correct across three consecutive intervention sessions. Similar to criterion on mastery of within-modal word learning, complete cross-modal generalization was defined as 80% for each direction of generalization. More broadly, to make decisions on whether each participant demonstrated successful cross-modal generalization of word learning, success in generalization for each direction was defined as meeting criterion on the majority of phases of intervention (two out of three phases).

### Interobserver agreement (IOA)

IOA was calculated for 33% of probe sessions and at least 33% of intervention sessions ($M = 40\%$, range: 34–50%) for each participant. IOA was conducted by an independent observer (another clinician) by watching probe or intervention sessions via videotape and recording the child’s responses to probes as either correct, incorrect, or no response. Words produced with articulation errors were marked “correct” as long as the response was intelligible. The observer’s scoring was then compared to the treating clinician’s record of the child’s scores. Total percentage agreement was calculated by the number of opportunities in which two observers agreed divided by the total number of opportunities for agreement. Total percentage agreement was above 93% for all reliability sessions.

### Procedural fidelity

Procedural fidelity was measured for at least 33% of baseline and probe sessions ($M = 56\%$, range: 44–63%) and at least 33% of intervention sessions ($M = 39\%$, range: 34–46%) for each participant. Procedural fidelity was conducted by an independent observer either during the session or via event recording while watching videotapes of sessions. The observer recorded the extent to which the intervention and probe phase trials were implemented as planned using a dichotomous system (implemented/not implemented). The following items were recorded for each trial: the clinician secured the child’s attention, the clinician delivered task directives correctly, the response interval was the correct length, the clinician praised testing behavior, an appropriate corrective feedback was given, and the intertrial interval was the correct length. Implementation accuracy was calculated across trials during interventions (20 trials) or across probes during probe session (48 probes). Procedural fidelity ranged from 93 to 100% across sessions for all participants.

### Results

Data from individual participants were graphed and analyzed at the participant level to answer all four research questions. All 10 participants successfully learned target vocabulary words to criterion (80%, across three consecutive sessions at the end of each intervention phase) in trained modalities, indicating that the combined storybook and play intervention was effective in teaching new vocabulary and that cross-modal generalization could be tested in all participants. However, substantial individual variability was observed in number of sessions needed to reach mastery criterion and pattern of word learning. The length of each intervention phase ranged from 6 to 34 sessions for each set of vocabulary trained to reach mastery criterion. Representative data from three participants are presented in Figures 1, 3, and 4. Plots of all other participant’s learning and testing data at baseline, during intervention, and follow-up probes are available in supplementary materials (Figures S1–S7). All participants’ performance on the cross-modal probes is presented in Table 2.

### Cross-modal generalization

Overall, 3 out of 10 children (Participant 4, 5, and 7; see Table 2) demonstrated successful cross-modal generalization of trained vocabulary for both expressive-to-receptive and receptive-to-expressive directions (i.e. 80% on cross-modal probes for at least two out of three intervention phases) similar to what is often seen in typical development. An example plot from one participant (Participant 4) who demonstrated successful cross-modal generalization for both directions is displayed in Figure 1.

For this participant, during both expressive-only and receptive-only intervention conditions, his demonstration of target words increased in level as compared with
the baseline probe conditions and had an overall accelerating trend for the majority of intervention phases. The percentage of correct identification remained at 0 or 25% for all target words prior to any intervention. For receptive-only trained vocabulary words, he reached criterion in seven, five, and seven sessions, respectively, for three phases. For expressive-only trained vocabulary words, he reached criterion in five, five, and eleven sessions, respectively, for three phases. He showed an immediate increase in the percentage of correct identification for both Phase 1 and Phase 2 with intervention. For Phase 3, though his performance remained low for the first four sessions, he demonstrated a steady increase since the 5th session and eventually reached criterion during the 11th session. His performance on the untrained control set of vocabulary remained at 0 or 25% throughout the probe conditions, which eliminated the internal threats of maturation and incidental learning.

Overall, cross-modal generalization was higher for the expressive-to-receptive direction than for the receptive-to-expressive direction. Nine out of ten children demonstrated successful cross-modal generalization on the expressive-to-receptive probes for the majority of phases (i.e. two out of three phases), with an average expressive-to-receptive generalization rate at 89% (ranging from 44 to 100%; Figure 2). In contrast, only three children (Participant 4, 5, and 7) reached mastery criterion for cross-modal generalizations on the receptive-

### Table 2. Participants' success in cross-modality probes during the probe sessions immediately following intervention.

| Participant | Percentage of tiers with successful cross-modality generalization |
|-------------|---------------------------------------------------------------|
|             | Expressive to receptive | Receptive to expressive |
| 1           | 100                   | 0                     |
| 2           | 100                   | 33                    |
| 3           | 0                     | 0                     |
| 4           | 100                   | 100                   |
| 5           | 67                    | 100                   |
| 6           | 100                   | 33                    |
| 7           | 100                   | 67                    |
| 8           | 67                    | 0                     |
| 9           | 67                    | 0                     |
| 10          | 67                    | 33                    |

![Figure 1](image1.png)

**Figure 1.** Vocabulary learning and cross-modal generalization of a participant who demonstrated successful cross-modal generalization in both directions.
to-expressive probes for the majority of intervention phases, with an average receptive-to-expressive generalization at 43% (ranging from 6 to 89%, see Figure 2). On average, participants demonstrated successful cross-modal generalization on four out of six words for expressive-only trained words but only one out of six words for receptive-only trained words.

An example plot from one participant (Participant 9) who demonstrated successful cross-modal generalization in the expressive-to-receptive direction is displayed in Figure 3. For this participant, his percentage of correct identification gradually and consistently increased after the intervention was introduced, although variability was observed for Phase 2 of the expressive-only condition and Phase 3 of the receptive-only condition. For the expressive-only condition, he reached criterion in 19, 14, and 22 sessions, respectively, for three phases. For receptive-only vocabulary words, he reached criterion in 9, 17, and 10 sessions, respectively. As shown in Figure 3, though this participant reached successful generalization criterion for two out of three phases for the expressive-to-receptive generalization (100, 83, and 50%), his performance on the receptive-to-expressive generalization remained low (33, 67,
and 33%). His performance on the control set of vocabulary remained low for the receptive-only condition, yet some variability (ranging from 0 to 50%) was observed for the control set of vocabulary for the expressive-only condition. However, because his percentage of accuracy was still under the mastery criterion, we concluded that the variability observed in the percentage of correct for the control vocabulary did not provide sufficient evidence for incidental learning. Five other participants (Participant 1, 2, 6, 8, and 10) demonstrated similar cross-modality generalization pattern only in the expressive-to-receptive direction. Finally, one participant did not demonstrate successful cross-modal generalization in either direction (Participant 3, Figure 4). For Phase 1 and Phase 2, with the introduction of intervention, his percentage of correct identification remained low for the first half of the intervention phase but increased with an accelerating trend during the second half of both phases. For Phase 3, significant variability was observed for the expressive condition, but he eventually reached criterion after 34 sessions. For the receptive-only condition for Phase 3, he showed an overall increasing trend immediately after the introduction of the intervention and maintained a high percentage of correct identification after reaching the criterion. However, he did not demonstrate successful cross-modal generalization in either direction even after he demonstrated mastery in the trained modality. His percentage of correct identification on the cross-modal probes ranged from 33 to 50% for the expressive-to-receptive generalization probes and from 0 to 50% for the receptive-to-expressive generalization probes.

Expressive and receptive word learning

Even though all children were successful in learning targeted vocabulary to mastery criterion in both expressive and receptive modalities, large variability was observed with regard to number of intervention sessions needed to reach mastery criterion and there were individual patterns of word learning. Total length of intervention ranged from 17 to 70 sessions ($M = 32$, $SD = 17$) for the expressive-only condition and ranged from 14 to 47 sessions ($M = 26$, $SD = 11$) for the receptive-only condition (Figure 5). Three participants (Participant 1, 6, and 7) mastered target vocabulary faster in the expressive-only condition on average compared to the receptive-only condition. Six participants (Participant 2, 3, 4, 5, 8, and 9) learned target vocabulary faster in the receptive condition. One participant (Participant 10) took an equal number of sessions to reach criterion for the two modalities. When participants’ vocabulary learning was
analyzed by intervention phase, two participants (Participant 2 and 5) consistently demonstrated faster learning in the receptive-only condition (3/3 replications of reaching criterion in the expressive-only condition before the receptive-only condition) for all three intervention phases. One participant (Participant 1) showed a consistent faster learning in expressive condition for all three intervention phases.

Measure of incidental (untreated) vocabulary learning

A fourth set served as an untreated control set to monitor possible incidental learning of vocabulary. The mastery criterion used for daily intervention probes and cross-modal probes was also used for the control set. One participant (Participant 2) demonstrated evidence of learning for one control word in the receptive condition. No other children demonstrated evidence of incidental learning to criterion in either modality. Therefore, these results demonstrate vocabulary intervention effects well above the observed levels of incidental word learning.

Discussion

The primary purpose of this study was to evaluate cross-modal generalization of vocabulary in children with ASD. Cross-modal generalization can be and is often assumed or “automatic” outcome in clinical practice. To be sure, rapid expressive-to-receptive and receptive-to-expressive generalization is common in TD children (Dollaghan, 1985; Fernandes, 2008). To systematically test this assumption in ASD, this study used an orthogonal test of expressive- and receptive-only vocabulary learning in which vocabulary targets were taught and probed independently in each modality. During intervention, all participants demonstrated within-modal learning for both expressive- and receptive-only conditions (30 replications across 10 participants). However, contrary to the assumption that vocabulary learning will be “automatically” generalized across modalities, results from this study indicate that cross-modal generalization at the word level is not automatic nor consistent in children with ASD, particularly in the receptive-to-expressive direction. Additionally, as a secondary analysis, we compared the efficiency of word learning in the expressive-only condition and the receptive-only condition to understand whether learning in one modality proceeds the other. Though all participants demonstrated learning in both modalities eventually, participants varied in terms of their learning pattern of the target vocabulary words in each modality.

Cross-modal generalization

In a broader sense, generalization refers to the process of transferring learned skills to a similar but not identical new context (Harris, 1975; Wolfe, Blankenship, & Rispoli, 2018). It is a ubiquitous challenge in clinical practice and is one of the most critical barriers to treatment success because learned skills in the clinical setting with trained materials may not readily generalize to functional language outcomes outside of the clinic context (de Marchena et al., 2015) Many children with intellectual disabilities showed poor performance on generalization tasks even when they mastered the materials or behaviors that were explicitly trained (Davis et al., 2016a, 2016b). For children with ASD, generalization has been known to be a particular challenge and past studies have consistently shown that children with ASD were less efficient in their generalization of learned skills in different domains including language and literacy, emotion recognition, and social skills (for a review, see Wass & Porayska-Pomsta, 2014).

In this study, we extend the current literature on generalization in children with ASD by examining a specific type of generalization in word learning: cross-modal generalization. Unlike children with typical language development, children with ASD in this study...
displayed two main patterns of cross-modal word generalization. First, children with ASD demonstrated more cross-modal generalization in the expressive-to-receptive direction than the receptive-to-expressive direction. Nine children demonstrated successful expressive-to-receptive generalization of trained words on at least two out of three intervention phases, yet only three children demonstrated successful generalization in the receptive-to-expressive direction. Second, even in the children who reached successful cross-modal generalization criteria, very few demonstrated consistent generalization across all three intervention phases. Consistent with results from Wynn and Smith (2003), these results provide a convergent picture of atypical vocabulary learning profile and weakness in cross-modal generalization in children with ASD, particularly from the receptive to expressive modality.

However, our results contrast with Wynn and Smith (2003) who found that children’s overall language ability had an impact on their levels of cross-modal generalization. In their study, the three participants who scored the highest on the standardized language assessment also demonstrated the highest level of cross-modal generalization. It is reasonable to speculate that a child’s language or cognitive ability could have an association with his or her ability to generalize across modalities or contexts. For instance, generalization requires the ability to relate newly encountered stimuli or context to past learning experiences (Rimland, 1964). Therefore, children with better working memory or short-term memory could be at an advantage for generalization tasks. Additionally, children with more advanced language skills could rely on their linguistic resources to draw meaningful connections between familiar and new contexts or materials, which could also contribute to greater success with generalization. However, we did not detect any pattern using visual analysis between children’s generalization ability and their language functioning, cognitive skills, or age in our study. The different findings in the association between cognitive or language skills and cross-modal generalization between this study and Wynn and Smith (2003) could arise from several sources including different sample distributions on these parameters. In general, single subject designs do not include systematic replication across multiple participant features so these kinds of analyses rely on categorical or binary distinctions. The hypothesis that cognitive or language abilities drive cross-modal generalization ability or broader generalization ability should be empirically tested in future studies with more extensive replication of these distinctions.

Additionally, our finding that children with ASD generalized more from the expressive modality to the receptive modality than the opposite direction inspired us to reflect more deeply about the relationship between expressive and receptive modalities of word learning and factors that may facilitate one direction of generalization over another. A task requirement analysis of cross-modal generalization suggests that expressive-to-receptive generalization requires a child to identify an item out of an array of four pictures and respond to a verbal prompt (i.e. “Where is the toucan?”) with a motor action (i.e. pointing), whereas receptive-to-expressive generalization requires a child to retrieve a label when presented with a picture and respond to a verbal prompt (i.e. “What is it?”) with a verbal answer when the individual item is presented.

The comparison of task requirements suggests that the task demand of verbal production of the item in the receptive-to-expressive generalization may result in greater complexity of the task than the task demand of pointing in the expressive-to-receptive generalization. Even though we attempted to control for participants’ oral-motor and speech production skills by only including words that a child can imitate as target vocabulary, we did not explicitly test participant’s phonological skills. Previous theoretical models on lexical learning and access have suggested that a speaker typically go through two stages to produce a word: lexical selection and phonological form encoding (Levelt, 2001; Levelt, Roelofs, & Meyer, 1999). During the lexical selection stage, the speaker selects the target lexical item from active lexical concepts in the mental lexicon. Once a lexical item is selected, the phonological form encoding system is activated to retrieve the sounds that correspond to the word that the speaker wants to produce.

It is possible that the additional step of retrieving phonological forms of the target vocabulary in the receptive-to-expressive generalization posed additional challenges for children with ASD in this study. However, evidence from previous studies on phonological skills of children with ASD, though limited, have generally suggested differences in phonological processing skills are not sufficient explanations of word learning difficulties in this population (Gladfelter & Goffman, 2017; Norbury, Griffiths, & Nation, 2010). In one study, children with ASD learned the phonological forms of new vocabulary and even outperformed TD peers when tested immediately after the word learning paradigm (Norbury et al., 2010). The authors interpreted their findings as indicative of enhanced attention to phonological forms in children with ASD and a “sound before meaning” strategy that they use to learn words. It is relevant to note that participants in the Norbury study were high-functioning children with ASD whereas our study included children with ASD with varying levels of cognitive skills. Further studies are needed to systematically examine...
the role that that phonological skills play in word learning in children with ASD.

Given the atypical receptive-expressive vocabulary profile in children with ASD (Davidson & Ellis Weismer, 2017; Hudry et al., 2010; Kover et al., 2013), we suspect that the particular difficulty in the receptive-to-expressive generalization observed in this study reflects a unique difficulty in receptive vocabulary learning or reduced receptive advantage in children with ASD rather than a difference in task demands across expressive and receptive vocabulary learning. Previous studies on the discrepant receptive-expressive vocabulary profile in children with ASD have suggested that children with ASD may demonstrate focused impairments in receptive vocabulary learning because they benefit less from the ambient linguistic input in the environment and may require more supportive contexts to learn new words (McDaniel, Yoder, Woynaroski, & Watson, 2018; Woynaroski et al., 2016). Specifically, McDaniel et al. (2018) proposed that deficits in attention toward the speaker in children with ASD may disrupt receptive vocabulary learning more than expressive vocabulary learning. In the same study, the authors provided evidence that reduced attention toward the speaker was associated with larger later receptive-expressive vocabulary size discrepancy in children with ASD. One possible explanation of the better expressive-to-receptive generalization in our study is that the expressive-only vocabulary learning condition provided a more supportive context for vocabulary learning and generalization for children with ASD than that of the receptive-only condition.

Additionally, based on this post hoc consideration of the receptive-expressive asynchrony, we speculate that the disparity in the extent of generalization across the two directions could also potentially be attributed to differences in teaching and prompting contexts across conditions. Even though the clinician provided two prompts for each target word during daily probes for both conditions, in the expressive condition the child was presented with one picture each time the clinician prompted the child to label the picture. However, for the receptive prompts in the receptive condition, the child was shown four pictures and was asked to point to the picture that matches the target word. Previous work on multisensory processing in ASD has found that children with ASD demonstrated deficits in perceiving the temporal relationship between cross-modal sensory inputs (i.e. audio and visual (A–V) inputs) and integrating information from multiple sensory modalities (Marco, Hinkley, Hill, & Nagarajan, 2011). Specifically, individuals with ASD have been reported to experience an extended or altered multisensory temporal binding window (Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011; Marco et al., 2011; Stevenson et al., 2014). Foss-Feig et al. (2010) showed that school-aged children with ASD exhibited a wider multisensory temporal binding window and were more likely to report asynchronous A–V stimuli as synchronous than participants with typical development. Deficits in A–V binding in individuals with ASD are not specific to low-level stimuli (i.e. auditory tones) but are also evident in tasks that involve simple or complex linguistic stimuli (Bebko, Weiss, Demark, & Gomez, 2006; Stevenson et al., 2014).

In our study, the expressive prompts in the expressive-only condition involved the visual cue and the auditory cue of the target item with simultaneous focus on one item (the target word). In contrast, the receptive prompts in the receptive-only condition always involve the visual information of three other items in addition to the visual cue and the auditory cue of the target item. Word learning is inherently multisensory in nature because young learners are expected to combine complementary information from A–V modalities when acquiring words. Thus, for children with ASD who may have variable temporal binding windows for A–V inputs and thus ambiguity in how these multisensory signals should be combined, the exclusive (and item-focused) binding between the A–V information of the target item in the expressive-only condition (i.e. one spoken word label and one referent) perhaps facilitated the mapping between the label and its referent by reducing the variability of stimuli and the complexity of cross-modal A–V multisensory integration. A hypothesis to address the observed greater expressive-to-receptive generalization is that the strengthened link between the label and its referent in the expressive-only condition contributed to a more consistent temporal binding of the auditory and visual features of the target word and facilitated cross-modal expressive-to-receptive generalization. Future studies are necessary to test this hypothesis and systematically examine the impact of potential deficits in multisensory integration on word learning in children with ASD.

**Potential autonomy for expressive and receptive modalities in ASD**

Another difference in the expressive and receptive prompts and teaching contexts could potentially explain the disparity in the extent of generalization across the two directions. In the expressive-only condition, the expressive prompts asked the child to name the target item, whereas in the receptive-only condition, the receptive prompts asked the child to point to a picture that matches the target word. Consequently, in the expressive condition, participants were expected to produce the labels and were only reinforced following their correct production, yet in the receptive condition,
participants were reinforced after they correctly identify the target picture.

A plausible explanation is that the prompts in the expressive-only condition provided children with ASD with opportunities to practice retrieving phonological forms, which further helped them develop more robust lexical representations for target words in the expressive-only condition as compared to the receptive-only condition. In contrast, in the receptive-only condition, it is possible that participants learned to point to the picture of the referent via associative learning between the spoken label in the prompt and the picture instead of developing mental representations of the target words that capture the referential relations between the spoken label, the picture, and the symbolic representation of the word. This explanation is consistent with existing theories and evidence suggesting that children with ASD rely on associative learning mechanisms to learn words and are more likely to develop partial knowledge of learned words (Baron-Cohen, Baldwin, & Crowson, 1997; Henderson, Powell, Gareth Gaskell, & Norbury, 2014; McDuffie, Yoder, & Stone, 2006; McGregor et al., 2012; Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff, & Tager-Flusberg, 2007; Preissler, 2008). To illustrate, young children with ASD have been previously found to associate spoken labels with pictures of the referent instead of real objects (Parish-Morris et al., 2007). Such findings suggest that even though an associative style of word learning may allow children to establish a connection between the spoken label and the picture, these connections may not be sufficiently robust to generalize a newly learned word label to a referent object, a different context, or to freely retrieve the label when prompted in a new modality.

However, it is also noteworthy that even though participants in this study demonstrated more expressive-to-receptive generalization, there were incidences in which a participant was able to answer within-modal expressive probes accurately yet was unable to comprehend the target items that they were able to produce expressively. A similar notion that children with ASD demonstrate word production without understanding words’ meanings has been previously proposed by Charman et al. (2003). As we integrate this finding with the observed difference in generalization patterns, it may be important to bear in mind that Arunachalam and Luyster (2016) suggested that children with ASD are indeed capable of establishing robust lexical representations but may need more input, time, or practice than TD children. In our study, even though cross-modal generalization in either direction was not automatic, participants did perform better on cross-modal generalization probes in the expressive-only condition than the receptive-only condition perhaps because the expressive prompts included opportunities to practice retrieving (and producing) newly learned words. We argue that these findings support the aforementioned proposal that input and output that include more learning opportunities could potentially enhance lexical representation in children with ASD and promote generalization. Granted, this study was not designed to test lexical representations in children with ASD, more research is needed to improve our understanding of lexical representations in this population and to shed light on how to provide input and prompts to optimize word learning and cross-modal generalization in children with ASD.

**Word learning efficiency and cross-modal generalization in ASD**

In addition, a secondary finding from this study suggests that unlike TD children whose receptive vocabulary is often considered to precede expressive vocabulary learning (Bornstein & Hendricks, 2012), children with ASD in this study varied in terms of their word learning patterns of the target words. Half of the participants reached mastery criterion for within-modal learning faster in the receptive-only intervention condition compared to the expressive-only condition. When participants’ learning was analyzed by intervention phase separately, one participant consistently demonstrated faster learning in the receptive condition across all three intervention phases and two children consistently reached criterion in the expressive-only intervention prior to the receptive-only modality. A closer examination of the learning data of these three children reveals that the difference in number of sessions to reach criterion across modalities was less than three sessions for each participant. These differences likely reflect individual variability instead of etiology-specific vocabulary learning patterns. Overall, participants’ learning data suggest that receptive- and expressive-only target vocabulary were learned in parallel and in close proximity to one another on a session-by-session basis.

**Clinical implications**

Findings from this study provide new insights into teaching expressive and receptive vocabulary for children with ASD. There has been a long-standing debate about when and in which sequence professionals should teach vocabulary to children with ASD. Results from this study show that more children demonstrated cross-modal generalization in the expressive-to-receptive direction than the receptive-to-expressive direction. When averaging across target word sets, participants demonstrated successful cross-modal generalization on
approximately four out of six words for expressively trained target words yet one out of six words for receptively trained words. These results suggest that targeting expressive vocabulary first with the goal of incidentally increasing receptive vocabulary may be more efficient than starting with the receptive modality. However, the finding that not all children demonstrated successful expressive-to-receptive cross-modal generalization indicates that teaching vocabulary in the expressive modality exclusively does not guarantee receptive understanding in this population. We recommend that practitioners periodically monitor children’s vocabulary learning in both modalities or set an explicit generalization goal to ensure complete learning of trained words.

Additionally, results from this study also highlight the need to probe receptive understanding of trained words if an expressive-first intervention approach is adopted. Specifically, some children in this study demonstrated 100% accuracy rate on within-modal expressive probes yet were unable to identify the target items that they were able to produce expressively. Such findings challenge the common clinical assumption that receptive knowledge of vocabulary typically precedes expressive production. It is thus essential to check children’s comprehension of trained words even after they were produced expressively.

**Limitations and future directions**

Results from this study should be interpreted within the context of some limitations. First, maturation and history are possible threats to internal validity in parallel treatments design studies because such a design usually requires a long time span to complete (Gast & Ledford, 2018). We assessed internal threats of maturation and history by examining participants’ performance on each vocabulary set in a probe session before instruction started. We also monitored incidental learning by including a fourth untrained control set of vocabulary. For all 10 participants, their performances on the probes before the introduction of each intervention condition remained unchanged (0–25%), which suggest that history and maturation were not occurring. For the fourth set of control vocabulary, we detected some evidence of incidental learning for two participants: for Participant 1, his performance on the within-modal probes for the untrained expressive control set ranged between 50 and 100% and for Participant 2, his performance on the within-modal probes for the receptive control set ranged between 75 and 100%. Their performance on the cross-modal probes remained unchanged. No other children demonstrated evidence of incidental learning to criterion in either modality. Overall, these results demonstrate intervention effects well above the observed level of incidental learning.

Another limitation of this study is that maintenance of participants learning was not systematically targeted or tested. Vocabulary instruction also took place within the clinical setting using toys that were closely linked to the theme of the book. It is plausible that teaching vocabulary in multiple contexts with an explicit treatment goal of maintaining learned words may enhance participants’ cross-modal generalization. Additional research is needed to determine the extent to which cross-modal generalization of word learning is malleable and explore effective treatment strategies that could improve cross-modal generalization.

In conclusion, this study is the first to indicate that children with ASD demonstrated inconsistent cross-modal generalization even after extensive within-modal training. Findings from this study warrant future studies to further understand (a) the relationship between modalities during word learning; (b) the optimal order of vocabulary instruction (i.e. simultaneous instruction, isolated instruction with one modality before another, expressive-before-receptive instruction, receptive-before-expressive instruction); and (c) potential interaction between language functioning, developmental level, and cognitive skills and children’s cross-modal generalization ability. For instance, does teaching vocabulary simultaneously in both expressive and receptive modalities help children with ASD develop more robust lexical representations for learned vocabulary? Does teaching expressive vocabulary globally translate to gains in receptive vocabulary globally? Answers to these questions will provide valuable insights into potential mechanisms of deficits in cross-modal generalization, identification of a subgroup of children with ASD who needs additional support with word acquisition and generalization, and new treatment strategies to improve treatment efficiency in this population.

**Acknowledgments**

The authors are grateful to the families who participated in this study and the clinicians and research assistants who assisted with the vocabulary interventions.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was completed with the support of award number R324A080143 to the final author from the
Institute on Educational Sciences (IES) and the support from the Scottish Rite Foundation of Nashville. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding organizations.

**Supplemental Material**

Supplemental material for this article is available online.

**References**

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research: Official Journal of the International Society for Autism Research*, 9(8), 810–828. doi:10.1002/aur.1590

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker’s direction of gaze strategy to crack the code of language? *Child Development*, 68, 48–57.

Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2004). Fast mapping of words and story recall by individuals with Down Syndrome. *Journal of Speech Language and Hearing Research*, 47(6), 1286. doi:10.1044/1092-4388(2004/097)

Bornstein, M. H., & Hendricks, C. (2012). Basic language comprehension and production in >100,000 young children from sixteen developing nations. *Journal of Child Language*, 39(4), 899–918. doi:10.1017/S0305000911000407

Bucher, B., & Keller, M. F. (1981). Transfer to productive labeling after training in comprehension: Effects of three training variables. *Analysis and Intervention in Developmental Disabilities, 1*(3–4), 315–331. doi:10.1016/0270-4684(81)90005-7

Camarata, S. M. (2014). Early identification and early intervention in autism spectrum disorders: Accurate and effective?. *International Journal of Speech-Language Pathology*, 16(1), 1–10. doi:10.1017/ISJ.2013.858773

Camarata, S. M., Nelson, K. E., Gillum, H., & Camarata, M. N. (2009). Incidental receptive language growth associated with expressive grammar intervention in SLI. *First Language*, 29(1), 51–63. doi:10.1177/0142723708098810

Carrow-Woolfolk, E. (1999). *TACL-3: Test for auditory comprehension of language*. Austin, TX: PRO-ED.

Charmian, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur communicative development inventory (infant form). *Journal of Child Language*, 30(1), 213–236. doi:10.1017/S0305000902005482

Davidson, M. M., & Ellis Weismer, S. (2017). A discrepancy in comprehension and production in early language development in ASD: Is it clinically relevant?. *47*, 2163–2175 doi:10.1007/s10803-017-3155-2

Davis, T. N., Camarata, S. M., & Camarata, M. N. (2016a). Cross-modal generalization of receptive and expressive vocabulary in children with Down Syndrome. *Journal of Down Syndrome & Chromosome Abnormalities*, 2(1), 1–9. doi:10.4172/2472-1115.1000105

Davis, T. N., Lancaster, H. S., & Camarata, S. M. (2016b). Expressive and receptive vocabulary learning in children with diverse disability typologies. *International Journal of Developmental Disabilities*, 62(2), 77–88. doi:10.1117/2047387715Y.000000010

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child and Adolescent Psychiatry*, 7(3), 131–136. doi:10.1007/s007870050058

de Marchena, A. B., Eistl, I. M., & Yerys, B. E. (2015). Brief report: Generalization weaknesses in verbally fluent children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 45(10), 3370–3376. doi:10.1007/s10803-015-2478-6

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, 95(3), 465–481. doi:10.1037/0022-0663.95.3.465

Dollaghan, C. (1985). Child meets word: “Fast mapping” in preschool. *Journal of Speech and Hearing Research*, 28, 449–454. doi:10.1044/jshr.2803.454

Doyle, P. M., Wolery, M., Gast, D. L., Ault, M. J., & Wiley, K. (1990). Comparison of constant time delay and the system of least prompts in teaching preschoolers with developmental delays. *Research in Developmental Disabilities*, 11(1), 1–22. doi:10.1016/0022-0663(90)90002-P

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Minneapolis, MN: Pearson Assessments.

Eldevik, S., Kazemi, E., & Elsky, G. (2016). Generalized use of past tense verbs in children with autism following a sufficient exemplar training procedure. *International Electronic Journal of Elementary Education*, 9(2 Special Issue), 485–498.

Ellis Weismer, S., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders*, 40(10), 1259–1273. doi:10.1007/s10803-010-0983-1

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2007). *MacArthur-Bates communicative development inventories: User’s guide and technical manual*. Baltimore, MD: Brookes.

Fernandes, K. (2008). *On the significance of speech: How infants discover symbols and structure*. ProQuest dissertations and theses. New York, NY: New York University.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., . . . Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203(2), 381–389. doi:10.1007/s00220-010-2240-4

Gast, D. L., & Ledford, J. R. (2018). Single case research methodology: Applications in special education and behavioral sciences (3rd ed.). New York, NY: Routledge.
Gast, D. L., & Wolery, M. (1988). Parallel treatments design: A nested single subject design for comparing instructional procedures. *Education and Treatment of Children, 11*(3), 270–285.

Gladfelter, A., & Goffman, L. (2017). Semantic richness and word learning in children with autism spectrum disorder. *Developmental Science, 21*(2), e12543. doi:10.1111/desc.12543

Goldstein, H. (2002). Communication intervention for children with Autism: A review of treatment efficacy. *Journal of Autism and Developmental Disorders, 32*(5), 373–396. doi:10.1023/A:1020589821992

Gray, S. (2003). Word-learning by preschoolers with specific language impairment. *Journal of Speech Language and Hearing Research, 46*(1), 56. doi:10.1044/1092-4388(2003/005)

Guess, D. (1969). A functional analysis of receptive language and productive speech: Acquisition of the plural morpheme. *Journal of Applied Behavior Analysis, 2*(1), 55–64. doi:10.1901/jaba.1969.2-55

Hampton, L. H., & Kaiser, A. P. (2016). Intervention effects on spoken-language outcomes for children with autism: A systematic review and meta-analysis. *Journal of Intellectual Disability Research, 60*(5), 444–463. doi:10.1111/jir.12283

Harris, S. L. (1975). Teaching language to nonverbal children with emphasis on problems of generalization. *Psychological Bulletin, 82*(4), 565–580. doi:10.1037/h0076903

Henderson, L., Powell, A., Gareth Gaskell, M., & Norbury, H. (1975). Teaching language to nonverbal children with specific language impairment. *Journal of Speech Language and Hearing Research, 46*(1), 56. doi:10.1044/1092-4388(2003/005)

Guess, D. (1969). A functional analysis of receptive language and productive speech: Acquisition of the plural morpheme. *Journal of Applied Behavior Analysis, 2*(1), 55–64. doi:10.1901/jaba.1969.2-55

Hampton, L. H., & Kaiser, A. P. (2016). Intervention effects on spoken-language outcomes for children with autism: A systematic review and meta-analysis. *Journal of Intellectual Disability Research, 60*(5), 444–463. doi:10.1111/jir.12283

Harris, S. L. (1975). Teaching language to nonverbal children with emphasis on problems of generalization. *Psychological Bulletin, 82*(4), 565–580. doi:10.1037/h0076903

Henderson, L., Powell, A., Gareth Gaskell, M., & Norbury, C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Developmental Science, 17*(6), 858–871. doi:10.1111/desc.12169

Herlihy, L., Knoch, K., Vibert, B., & Fein, D. (2015). Parents’ first concerns about toddlers with autism spectrum disorder: Effect of sibling status. *Autism, 19*(1), 20–28. doi:10.1177/1362361315509731

Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder – A follow-up comparison in early adult life II: Social, behavioural, and psychiatric outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(5), 561–578. doi:10.1017/S002204960005006

Hudry, K., Chandler, S., Bedford, R., Pasco, G., Gliga, T., Elsabbagh, M., ... Charman, T. (2014). Early language profiles in infants at high-risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(1), 154–167. doi:10.1007/s10803-013-1861-4

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., ... PACT Consortium. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders, 45*(6), 681–690. doi:10.3109/13688280903461493

Hus, V., Pickles, A., Cook, E. H., Risi, S., & Lord, C. (2007). Using the autism diagnostic interview-revised to increase phenotypic homogeneity in genetic studies of autism. *Biological Psychiatry, 61*(4), 438–448. doi:10.1016/j.biopsych.2006.08.044

Jones, C. D., & Schwartz, I. S. (2004). Siblings, peers, and adults: Differential effects of models for children with autism. *Tecce, 244*, 187–198.

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. doi:10.1080/01690960042000058

Kover, S. T., McDuffie, A. S., Hagerman, R. J., & Abbeduto, L. (2013). Receptive vocabulary in boys with autism spectrum disorder: Cross-sectional developmental trajectories. *Journal of Autism and Developmental Disorders, 43*(11), 2696–2709. doi:10.1007/s10803-013-1823-x

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., & Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in Integrative Neuroscience, 4*(January), 1–11. doi:10.3389/fnint.2010.00129

Levett, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences, 98*(23), 13464–13471. doi:10.1073/pnas.231459498

Levett, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences, 22*, 1–75. doi:10.1017/S0140525X99001776

Lord, C., Risi, S., Lambrech, L., Cook, E. H. J., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). The autism diagnostic schedule – Generic: A standard measures of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. doi:10.1021/A300592401947

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

Lovaa, O. I., Koegel, R., Simmons, J. Q., & Long, J. S. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis, 6*(1), 1310815. doi:10.1901/jaba.1973.6-131

Lyyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(8), 1426–1438. doi:10.1007/s10803-007-0510-1

Lyyster, R. J., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the Macarthur-Bates communicative development inventory (CDI). *Journal of Child Language, 34*(3), 623–654. doi:10.1017/S0305000907008094

Lyyster, R. J., & Lord, C. (2009). Word learning in children with autism spectrum disorders. *Developmental Psychology, 45*(6), 1774–1786. doi:10.1037/a0016223

Marco, E. J., Hinkley, L. B. N., Hill, S. S., & Nagarajan, S. S. (2011). Sensory processing in Autism: A review of neurophysiologic findings. *Pediatric Research, 69*(5 Pt 2), 48R–54R. doi:10.1203/PDR.0b013e3182130c54

Martin, N. A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test-4 (EOVPVT-4).* San Antonio, TX: Academic Therapy Publications.
McDaniel, J., Yoder, P., Woynaroski, T., & Watson, L. R. (2018). Predicting receptive-expressive vocabulary discrepancies in preschool children with autism spectrum disorder. *Journal of Speech Language and Hearing Research, 61*(6), 1426. doi:10.1044/2018_JSLHR-L-17-0101

McDuffie, A. S., Yoder, P. J., & Stone, W. L. (2006). Labels increase attention to novel objects in children with autism and comprehension-matched children with typical development. *Autism, 10*(3), 288–301. doi:10.1177/1362361306063287

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders, 42*(1), 35–47. doi:10.1007/s10803-011-1210-4

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., … Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics, 27*(2 Suppl. 2): S69–S78. doi:00004703-200604002-00004 [pii]

Norbury, C. F., Griffiths, H., & Nation, K. (2010). Sound before meaning: Word learning in autistic disorders. *Neuropsychologia, 48*(14), 4012–4019. doi:10.1016/j.neuropsychologia.2010.10.015

Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development, 78*(4), 1265–1287. doi:10.1111/j.1467-8624.2007.01065.x

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. doi:10.1002/aur.12

Pellecchia, M., & Hineline, P. N. (2007). Generalization of mands in children with autism from adults to peers. *The Behavior Analyst Today, 8*(4), 483–491. doi:10.1037/h0100634

Petursdottir, A. I., & Carr, J. E. (2011). A review of recommendations for sequencing receptive and expressive language instruction. *Journal of Applied Behavior Analysis, 44*(4), 859–876. doi:10.1901/jaba.2011.44-859

Pickles, A., Anderson, D. K., & Lord, C. (2014). Heterogeneity and plasticity in language: A 17-year follow-up of children referred early for possible autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 55*(12), 1354–1362. doi:10.1111/jcpp.12269

Preissler, A. M. (2008). Associative learning of pictures and words by low-functioning children with autism. *Autism, 12*(3), 231–248. doi:10.1177/1362361307088753

Rimland, B. (1964). *Infantile Autism*. East Norwalk, CT, US: Appleton-Century-Crofts.

Roid, G. H., & Miller, L. J. (2011). *Leiter international performance scale-revised (Leiter-R)*. Madrid, Spain: Psymtec.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., … Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience, 34*(3), 691–697. doi:10.1523/JNEUROSCI.3615-13.2014

Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner’s analysis of verbal behavior for children with autism. *Behavior Modification, 25*(5), 698–724. doi:10.1177/014544501255003

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

Taylor, B., & McDonough, K. (1996). Selecting teaching programs. In C. Maurice, G. Green & S. C. Luce (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 63–177). Austin, TX: PRO-ED.

Tek, S., & Naigles, L. (2017). The shape bias as a word-learning principle: Lessons from and for autism spectrum disorder. *Translational Issues in Psychological Science, 3*(1), 94–103. doi:10.1037/tps0000104

Wass, S. V., & Porayska-Pomsta, K. (2014). The uses of cognitive training technologies in the treatment of autism spectrum disorders. *Autism, 18*(8), 851–871. doi:10.1177/1362361314499827

Wolery, M. (2013). A commentary: Single-case design technical document of the what works clearinghouse. *Remedial and Special Education, 34*(1), 39–43. doi:10.1177/0741932512468038

Wolery, M., Ault, M. J., & Doyle, P. M. (1992a). *Teaching students with moderate to severe disabilities: Use of response prompting strategies*. White Plains, NY: Longman Publishing Group.

Wolery, M., Holcombe, A., Cybriwsky, C., Doyle, P. M., Schuster, J. W., Ault, M. J., … Gast, D. (1992b). Constant time delay with discrete responses: A review of effectiveness and demographic, procedural, and methodological parameters. *Research in Developmental Disabilities, 13*(3), 239–266.

Wolfle, K., Blankenship, A., & Rispoli, M. (2018). Generalization of skills acquired in language for learning by young children with autism spectrum disorder. *Journal of Developmental and Physical Disabilities, 30*(1), 1–16. doi:10.1007/s10882-017-9572-y

Woynaroski, T., Yoder, P. J., & Watson, L. R. (2016). Atypical cross-modal profiles and longitudinal associations between vocabulary scores in initially minimally verbal children with ASD. *Autism Research, 9*(2), 301–310. doi:10.1002/aur.1516

Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions, 18*(4), 245–266. doi:10.1002/bin.142

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1992). *PLS-3: Preschool Language Scale-3*. San Antonio, TX: Psychological Corporation.