



Children with Autism Spectrum Disorder May Learn from Caregiver Verb Input Better in Certain Engagement States

Madison Cloud Crandall¹ · Kristen Bottema-Beutel² · Jena McDaniel³ · Linda R. Watson⁴ · Paul J. Yoder¹

Published online: 9 May 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The relation between caregiver follow-in utterances with verbs presented in different states of dyadic engagement and later child expressive verb vocabulary in children with autism spectrum disorder (ASD) was examined in 29 toddlers with ASD and their caregivers. Caregiver verb input in follow-in utterances presented during higher order supported joint engagement (HSJE) accounted for a significant, large amount of variance in later child verb vocabulary; $R^2 = .26$. This relation remained significant when controlling for early verb vocabulary or verb input in lower support engagement states. Other types of talk in follow-in utterances in HSJE did not correlate with later verb vocabulary. These findings are an important step towards identifying interactional contexts that facilitate verb learning in children with ASD.

Keywords Autism spectrum disorder · Verbs · Caregiver input · Engagement state · Language

The development of spoken language by the preschool years has been identified as an important predictor of long-term adaptive outcomes for individuals with autism spectrum disorder (ASD; e.g., Howlin et al. 2004; Tager-Flusberg et al. 2009). For all children, including those with ASD who also have language learning difficulties, an important aspect of language development is the generative use of grammatical utterances (Eigsti et al. 2007). An important building block for forming grammatical utterances is a strong expressive verb vocabulary. Verbs facilitate the formation of early sentences by enabling communication about event structures and relations between objects (Brandone et al. 2007). At least in children with typical development (TD), early verb

lexical diversity predicts later grammar outcomes above and beyond early noun lexical diversity (Hadley et al. 2016).

Theory suggests that many verbs may be more challenging to acquire than nouns that are object labels (Gleitman et al. 2005; Golinkoff and Hirsh-Pasek 2006). Verbs tend to comprise a smaller proportion of total vocabulary than object nouns in young children with TD. This finding, which has been replicated across multiple languages, including American English (Bornstein et al. 2004; Golinkoff and Hirsh-Pasek 2006), may be a result of traits inherent to verbs as a general word class. Relative to nouns that label objects, verb referents are often more abstract or transient, more relational, and reflect more generalized classes of events (Golinkoff and Hirsh-Pasek 2006; Golinkoff and Hirsh-Pasek 2008; Johnson and de Villiers 2009; Lidz 2006).

Providing adequate contextual support for verb learning, while also minimizing processing and social demands, is important for children with ASD. At least early in language development, much of language learning occurs in the context of social interactions with an adult partner. While morphosyntactic, social, interactional, and perceptual cues all may be present in the interaction environment, children place different emphasis on each cue type in different stages of development (Maguire et al. 2006). Morphosyntactic cues certainly are helpful for verb learning, but social cues are likely to be more useful earlier in development (Maguire et al. 2006). However, because

✉ Madison Cloud Crandall
madison.l.cloud@vanderbilt.edu

¹ Special Education Department, Vanderbilt University, 110 Magnolia Circle, OMC 318, Nashville, TN 37212, USA

² Lynch School of Education, Boston College, 140 Commonwealth Ave, Chestnut Hill, MA 02467, USA

³ Department of Hearing and Speech Sciences, Vanderbilt University, 1215 21st Avenue South, 8310 MCE, South Tower, Nashville, TN 37232, USA

⁴ Division of Speech and Hearing Sciences, University of North Carolina at Chapel Hill, CB# 7190, Chapel Hill, NC 27599-7190, USA

of social communication deficits, children with ASD may miss some of these cues. The demands of many social interactions may exceed the cognitive resources of early language learners with ASD. Exceeding such cognitive resources may result in children with ASD needing multiple opportunities to learn new words. Similarly, many populations of children with language learning difficulties require more exposures to the same word-referent associations for optimal learning to occur (McGregor et al. 2013; Plante et al. 2011).

Children with ASD experience challenges associated with social attention deficits. Despite these deficits, children with ASD may be most likely to learn new words when (a) they are presented in the context of follow-in utterances, and (b) the caregiver and child are in a particular type of engagement state, which we will later refer to as higher order supported joint engagement (HSJE; Adamson et al. 2009; Bottema-Beutel et al. 2014).

The Importance of Follow-In Utterances

One way adults may increase the level of support and decrease social demands for verb learning for children with ASD is to tailor their talk to relate to the child's momentary focus of attention. Adult utterances that map onto a child's attentional focus are called *follow-in utterances*. Follow-in utterances do not require children to shift their attention from their current focus to the referent of the adult's utterance. Therefore, follow-in utterances might increase the probability that a child with ASD will accurately identify the meaning of the adult's utterance (McDuffie and Yoder 2010).

In a sample of children with ASD, caregiver use of follow-in utterances predicted later child language development more strongly than other types of caregiver utterances (McDuffie and Yoder 2010). Likewise, in a sample of initially preverbal children with ASD, the frequency of caregiver follow-in utterances was one of the four predictors of expressive language remaining significant after controlling for eight other known predictors of language development (Yoder et al. 2015).

Although most of the work on follow-in utterances in children with ASD focuses on total vocabulary, which may primarily include nouns, in children with TD, the frequency of follow-in verb input related positively to child verb vocabulary several months later (Tomasello and Kruger 1992). In contrast, the frequency of verbs presented outside of follow-in utterances *negatively* related to later child verb vocabulary (Tomasello and Kruger 1992). A positive relation between the frequency of follow-in verb input and later child verb vocabulary was also found in a sample of young children with ASD (Crandall et al., in press).

The Importance of Engagement State

Even the best crafted input may be less processable for children with ASD if it is presented in engagement states that do not take into consideration attentional and interactional demands. HSJE may be one engagement state that maximizes processability of input. In HSJE, the child and adult are acting on the same object set, the adult's actions influence the child's actions in some manner, and the child shows evidence that he or she is responsive to the adult's involvement in the interaction by reciprocally engaging (Bottema-Beutel et al. 2014). Within HSJE, the child does not use two- or three-point gaze shifts between the adult and object (i.e., coordinated gaze shifting), a skill that is particularly difficult for many children with ASD (Dawson et al. 2004). Instead, the child uses imitation, compliance, turn-taking, and/or intentional communication not involving coordinated gaze shifts to demonstrate evidence that he is responsive to the adult's role in the object-oriented activity and to continue the interaction (Bottema-Beutel et al. 2014).

Additionally, HSJE indicates more than momentary coordinated attention to object and adult. The interaction is sustained for a threshold period of time. For example, envision a child pushing a train along a track the adult is building, and then he stops pushing and waits for the adult to place another track piece before continuing to push the train. The adult places another track piece. The child commences pushing the train on the newly placed track. By stopping and waiting for the adult to take her "turn", the child demonstrates clear attention and responsivity to the adult's role as a partner in the interaction.

In contrast, in lower order supported joint engagement (LSJE) the child's play is influenced by the adult, but the child shows little to no explicit behavioral evidence of responsivity to the adult's involvement and does not reciprocally respond to the adult's interactional overtures (Bottema-Beutel et al. 2014). For example, the child might push the train along a track as the adult lays down additional pieces of the track, but the child does not wait for the adult's "turn" before pushing the train further. Similar to LSJE, Object Engagement (OE) is characterized by a lower level of support for processing the caregiver's talk than HSJE. The adult's actions do not exhibit an influence on the child's play in the OE state. Evidence of child responsivity to an adult's participation in object-centered activities may be important for children with ASD because it may signal brief intervals in which the child is processing the adult's input.

Evidence suggests that caregiver follow-in utterances in LSJE may be less likely to produce an association between the words in the caregiver's utterances and their

referents (i.e., be less processable) to children with ASD than follow-in utterances in HSJE. In a sample of children with ASD, time in HSJE was predictive of later language even when controlling for time in LSJE and in OE (Bottema-Beutel et al. 2014). A sub-sample of this sample of children will be used to test the present study's research questions.

Verb Learning in HSJE

HSJE may be a facilitative engagement state for verb learning. The reciprocal, active, and sustained participation of the child with his or her caregiver around an activity involving an object may give the child multiple opportunities to be exposed to (a) processable verb input and (b) highly-salient, active experience of the verb referent. Thus, these repeated, high-quality opportunities may increase the probability that the child will understand the meaning of the verb input. This is important, as understanding verbs provides the semantic basis for verb use. HSJE may also provide opportunities for the child to learn the interactive contexts in which verbs are used expressively. For example, a child looks at his caregiver pushing a train. After a few seconds, the caregiver enthusiastically says, "You push one too." Seeing the adult enjoying the activity, the child might then begin to push another train behind his or her caregiver's train. Not only is the verb input about an action referent the child is already attending to (i.e., pushing), the child is also showing responsivity to the adult's suggested action by complying with the adult's direction to push the train. The child can also connect the caregiver's talk with specific pragmatic functions—for example, to invite a partner to engage in an activity. Finally, the episode is sustained for a sufficient period of time to allow the child to slowly process the meaning and potential pragmatic functions of words. Understanding the pragmatic function that utterances with verb constructions can be used for (e.g., to direct a partner to produce an action) may support the child in using them expressively in the future.

The above example is offered in contrast to the child independently completing an action while a caregiver narrates his or her play, which is considered insufficient support for verb learning for many children with ASD. When the caregiver narrates the child's action, he or she provides verb input in follow-in utterances embedded in OE. For example, the child pushes the train. The adult watches the child play and says, "You are pushing the train." In this example, the child does not show any explicit evidence that he or she is aware of the adult or responsive to their linguistic input. Given the social communication and interactional difficulties in children with ASD, and the referential challenges verbs present, the child may not process the verb input. Further, in this example the child is not engaged

in a reciprocal interaction that would support an understanding of the interactive functions of verb phrases.

Research Questions

The primary research question in this study is "Does HSJE with caregiver follow-in verb input (HSJE + FV) predict later child expressive verb vocabulary in children with ASD?" The current study provides an initial examination of whether the engagement state in which follow-in verb input occurs matters. Similar to many children with language learning problems, children with ASD with language learning difficulties may need multiple exposures of the same verb in processable contexts to learn it. Additionally, amount of follow-in verb input is highly intercorrelated with diversity of follow-in verb input in naturalistic interactions (Crandall et al., in press). Thus, in the current study, we focused on *amount* of follow-in verb input in various engagement states. Children with ASD may benefit from multiple exposures to the same verb when provided in HSJE.

Additionally, we addressed three secondary research questions aimed at ruling out possible alternative explanations for observed significant findings. The first alternative explanation to be addressed is one that is common in studies evaluating relations between caregiver talk and child language: early variation in child language is stable over time and covaries with caregiver input (Yoder and Kaiser 1989). To rule out this explanation, we asked, "Does HSJE + FV predict later child expressive verb vocabulary after controlling for earlier child expressive verb vocabulary?" Second, previous findings have shown that the quantity of follow-in verb input predicted later verb vocabulary in children with ASD (Crandall et al., in press). It is possible that providing follow-in verb input in general, not follow-in verb input *in HSJE* specifically, is related to later child verb vocabulary. To address this potential alternative explanation, we asked, "Does HSJE + FV predict later child expressive verb vocabulary after controlling for verb input in follow-in utterances in LSJE or OE states?" Third, it is possible that providing linguistic input in follow-in utterances while in HSJE in general, not *verb* input specifically, is related to later child verb vocabulary. To test this potential explanation, we asked, "Do non-verb follow-in utterances provided in HSJE predict later child expressive verb vocabulary?" with an expectation that it would not.

Methods

Study Design

This study utilized a longitudinal correlational design. Three measurement points were used. The study spanned twelve

months in total. Time 1 and Time 2 were separated by four months. Time 2 and Time 3 were separated by eight months.

Participants

Participants were selected from a larger longitudinal correlational study aimed at identifying predictors of useful speech in children with ASD aged 24–48 months (Yoder et al. 2015). Child participants all were diagnosed with ASD according to criteria in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition-Text Revision (American Psychiatric Association 2000). Licensed or research-reliable clinicians confirmed these diagnoses at entry into the larger study using the Autism Diagnostic Observation Schedule, second edition, Module One (ADOS-2; Lord et al. 2012). The Vanderbilt University and University of North Carolina at Chapel Hill Institutional Review Boards approved study procedures. Caregivers provided written consent for participants.

Participants were selected from the larger study sample based on two criteria. First, participants had to have a caregiver report of at least 50 words produced on the MacArthur-Bates Communicative Development Inventory, Words and Gestures form, at the final measurement point (MCDI; Fenson et al. 2007). This criterion was selected to increase the likelihood that most participants would have at least some verbs in their expressive vocabularies by the final measurement point. The mean total expressive vocabulary

at the final time point for the selected participants was 153 words ($SD = 88.52$). Second, participants had to have existing data at the time point used to obtain caregiver input in engagement state variables (i.e. the current study's Time 1). Twenty-nine participants and their parents were identified for inclusion in this study. At Time 1, children were on average 38.69 months old ($SD = 7.43$). At Time 2, children were an average of 42.48 months old ($SD = 7.48$). At Time 3, children were an average of 50.52 months old ($SD = 7.54$). Table 1 provides descriptive data for the participating children and their parents.

Measures

For the current study, we examined child verb vocabulary measures from MCDI reports administered at Times 2 and 3 (Fenson et al. 2007). Time 2, not Time 1, was selected as the “early” time point at which child expressive verb vocabulary was controlled because there was more variance in child verb vocabulary at Time 2 than at Time 1. Too little variance in child verb vocabulary at Time 1 would attenuate associations with caregiver input and Time 3 child expressive verb vocabulary, thus reducing the importance of controlling for early verb vocabulary. The Time 2 measurement point thus afforded a stronger test of early verb vocabulary as an alternative explanation for observed relations than did controlling for verb vocabulary at Time 1. Caregiver verb input in engagement state variables were extracted from

Table 1 Sample descriptive statistics

Variable	Mean	<i>SD</i>
Demographics		
Ratio female: male	7:22	
ADOS Total Algorithm Score	21.43	4.82
At larger study start ^a		
MSEL mental age	14.32	4.80
MSEL developmental quotient	43.00	15.00
MSEL expressive age equivalency (months)	10.04	4.19
MSEL receptive age equivalency (months)	7.12	6.90
Primary parent formal education	Some college	– 1 SD = high school graduate, + 1 SD = college graduate
At present study time 1		
Chronological age (months)	38.69	7.43
MCDI number of words said	26.00	22.68
MCDI number of words understood	137.61	104.26

ADOS Autism Diagnostic Observation Schedule, second edition, Module One (Lord et al. 2012); MSEL Mullen Scales of Early Learning (Mullen 1995)

Developmental quotient = mental age/chronological age \times 100; MCDI MacArthur Bates communicative development inventory, words and gestures form (Fenson et al. 2007)

^aFour months prior to present study's time 1

15-min parent–child free play sessions collected at Time 1. Although the media records of the sessions for the current study had been collected in the original study, these caregiver–child sessions were re-coded for our variables of interest. See Table 2 for variables derived from these procedures and for descriptive statistics.

Child Expressive Verb Vocabulary

The MCDI Words and Gestures form, a parent report vocabulary checklist, was used to measure total child expressive verb vocabulary (Fenson et al. 2007). The number of words reported as said (and understood) in the action word section of the form was used as an estimate of child expressive verb vocabulary.

Caregiver’s Follow-In Verb Input in Different Engagement States

Fifteen-minute parent–child free play sessions were video-recorded and coded to estimate the quantity of follow-in verb input in different states of engagement. For the free play sessions, a standard set of toys was provided. Caregivers were instructed to play as they normally would with their child at home.

In the larger study’s primary data collection period, videos were coded for the presence of caregiver follow-in utterances. Coding was conducted using a 5-s partial interval behavior sampling method using Procoder software (Tapp

and Walden 1993). Intervals were considered codeable based on two criteria: (a) both the child and adult were visible and (b) the child was not engaging in behavior requiring redirection or soothing (e.g., pulling out items from a mother’s purse or crying uncontrollably). A follow-in utterance was defined as a caregiver utterance that referenced the object and/or action the child was attending to (Yoder et al. 2015). These utterances could be directives or comments. For the current study, the intervals with follow-in utterances were then coded for the presence of a verb to yield an estimate of quantity of caregiver follow-in verb input (FV).

Engagement states were coded using a timed-event sampling method using Procoder software (Tapp, 2003). Start and stop times for each state were recorded. States had to last for at least 3 s to be coded. Although other states were coded, three states were analyzed for the current study (defined in Table 3).

Only HSJE, LSJE, and OE codes were of interest for the present study as these were the most common states in young children with ASD that are aligned with the test of the study’s motivating theory. Taken together, these three states account for a substantial proportion of total observation time. Participants spent, on average, 56% of the 15-min observation in OE ($SD = .19$), 10% in LSJE ($SD = .08$), and 6% in HSJE ($SD = .06$).

To quantify the number of intervals containing follow-in verb input that co-occurred within engagement states of interest, we used the Procoder Merger program (Tapp 2013). For each participant, the Merger program was

Table 2 Variable descriptive information

Construct	Procedure (times)	Variable metric	Role (aim)	Mean (<i>SD</i>)
Child expressive verb vocabulary	MCDI (Times 2 and 3)	Total action words reported as said (and understood)	Covariate (Time 2) Criterion variable (Time 3)	Time 2: 5.14 (5.63) Time 3: 17.79 (14.40)
Follow-in verb input in HSJE (HSJE + FV)	Parent–child free play (Time 1)	Count of codeable intervals in the session that were coded as (a) HSJE and (b) follow-in caregiver utterance with at least one verb	Predictor	4.76 (7.59)
Follow-in verb input in lower support engagement states (LSE + FV)	Parent–child free play (Time 1)	Count of codeable intervals in the session that were coded as (a) LSJE or OE and (b) follow-in caregiver utterances with at least one verb	Covariate	28.59 (16.52)
Follow-in nonverb input in HSJE	Parent–child free play (Time 1)	Count of codeable intervals in the session that were coded as (a) HSJE and (b) follow-in caregiver utterance that does not include a verb	Predictor	1.14 (2.01)

MCDI MacArthur Bates communicative development inventory, Words and Gestures form (Fenson et al. 2007); *HSJE* Higher order supported joint engagement, *LSE* lower support engagement states, which include lower order supported joint engagement (LSJE) and object engagement (OE)

Table 3 Engagement state definitions

State	Definition
Higher order supported joint engagement (HSJE)	Within a supported joint engagement framework, the child reciprocates the adult's actions or collaborates with the adult by turn taking sequences, imitating the other's action, the child complying with caregiver's verbal commands, and the child using heightened affect when the caregiver performs an action explicitly meant to elicit affect from the child
Lower order supported joint engagement (LSJE)	Within a supported joint engagement framework, the child engages in actions on objects but does not show explicit acknowledgment of the adult
Object engagement (OE)	The child manipulates toys independently. Caregiver may be an onlooker or provide narration to the child's play, but any caregiver actions do not affect the child's play with toys

Lower support engagement states include LSJE and OE

used to combine (a) the file coded with interval sampling (i.e., the ones coded for follow-in verbs) and (b) the file coded with timed-event sampling (i.e., the ones coded for engagement states). To convert timed-event data to interval data, the program assigns an engagement code to each interval based on the engagement state with the longest duration in that interval. Thus, each interval was coded for the presence of a follow-in verb and for engagement state. These files produced variables estimating quantity of caregiver follow-in verb input in the three key engagement states: (a) HSJE (HSJE + FV), (b) LSJE (LSJE + FV), and (c) OE (OE + FV). LSJE and OE were conceptualized as “lower” support states of engagement than HSJE and were considered similar in regards to proposed processability of the input provided. Consequently, the number of intervals with LSJE + FV and number of intervals with OE + FV were summed to create a combined category called “lower support state” verb input (LSE + FV). Non-verb follow-in utterances in HSJE was calculated by subtracting the number of intervals coded as HSJE with follow-in verbs from the total number of intervals coded as HSJE with a follow-in utterance. See Table 3 for descriptive information for these variables.

Reliability

Twenty percent of all parent–child free play sessions were randomly selected to be coded by a trained secondary independent coder. The primary coder was blind to which sessions were chosen for reliability checks. An intra-class correlation coefficient (ICC) value was calculated for each follow-in utterance variable and each engagement state variable extracted from the parent–child free play sessions. SPSS statistical software was used to calculate the ICC values using the absolute agreement and participant and observer as random factors. The reliability threshold of .70 was selected based on convention (Koo and Li 2016). The ICC values exceeded .70 for all variables (mean = .84).

Results

Data Analysis

Generalized linear models using maximum likelihood estimates of the coefficients and robust standard errors were used to test the relations of interest. These analyses did not indicate undue influence from outlier data points and are robust to violations of assumptions of normally distributed residuals and homoscedasticity (Erceg-Hurn and Mirosevich 2008).

Does Follow-In Verb Input Produced Within HSJE (HSJE + FV) Predict Expressive Verb Vocabulary in Children with ASD?

HSJE + FV significantly predicted later child expressive verb vocabulary. This association was large (see Model 1, Table 4).

Secondary Analyses

Does HSJE + FV Predict Later Child Expressive Verb Vocabulary After Controlling for Earlier Child Expressive Verb Vocabulary?

In a regression model with both earlier child expressive verb vocabulary and HSJE + FV as predictors and later child expressive verb vocabulary as a criterion variable, HSJE + FV remained a significant predictor. Additionally, the value-added relation of early caregiver HSJE + FV to later child expressive verb vocabulary was large after controlling for earlier child expressive verb vocabulary (see Model 2, Table 4).

Does HSJE + FV Predict Later Child Expressive Verb Vocabulary After Controlling for Follow-In Verb Input Produced Within LSE States?

In a regression model with both LSE + FV and HSJE + FV as predictors and later child expressive verb vocabulary as a

Table 4 Unstandardized regression coefficients, standard errors, confidence intervals for primary and secondary analyses

	Model 1			Model 2			Model 3			Model 4				
	<i>b</i>	SE <i>b</i>	95% CI											
HSJE + FV ^a	.76*	.28	[.22, 1.31]	.61*	.23	[.17, 1.05]	.73*	.30	[.15, 1.31]					
Early child expressive verb vocabulary				1.40**	.35	[.71, 2.07]								
LSE + FV ^b							.06	.12	[−.18, .30]					
Non-verb follow-in utterances in HSJE										−.14	1.15	[−2.39, 2.11]		
Total R ²				.26										
R ² change with addition of HSJE + FV					.25									

CI Confidence interval

^aFollow-in verb input in high-level supported joint engagement

^bFollow-in verb input in lower support states (i.e., lower order supported joint engagement or object engagement)

p* < .05; *p* < .005

criterion variable, HSJE + FV remained a significant predictor. Additionally, the value-added association of caregiver follow-in verb input in HSJE and later child expressive verb vocabulary was large after controlling for follow-in verbs in LSE states. Finally, follow-in verb input in LSE states was not a significant predictor of later child expressive verb vocabulary after controlling for follow-in verbs in HSJE (see Model 3, Table 4).

Do Non-Verb Follow-In Utterances Provided Within HSJE Predict Later Child Expressive Verb Vocabulary?

As predicted, non-verb follow-in utterances in HSJE were not significantly associated with later child expressive verb vocabulary, total *R*² = .0005, *ns* (see Model 4, Table 4). This test provided discriminant validity information in the sense that this relation was expected to be non-significant and small.

Discussion

This exploratory study identified a significant, positive, and large relation between early quantity of follow-in verbs in HSJE and later child expressive verb vocabulary. This relation remained significant when controlling for two potentially confounding covarying variables: earlier child expressive verb vocabulary and follow-in verb input presented within other states of engagement. Additionally, the quantity of non-verb follow-in utterances in HSJE was not associated with later expressive verb vocabulary. These findings support the specificity of the observed relation; it is quantity of follow-in *verb* input (not follow-in utterances that do not contain verbs) in *HSJE* (not in states with less support for processing input) that is related to later child expressive verb vocabulary.

Contributions to the Literature

To our knowledge, our findings are the first to confirm a prediction that follow-in verb input presented within one engagement state is differentially associated with later child expressive verb vocabulary in children with ASD. Our results extend the extant literature on the facilitation of expressive verb vocabulary in children with ASD in four ways.

First, the current study’s examination of *unmanipulated* interactions extend past experimental findings showing that children with ASD are sensitive to certain qualities of linguistic input that affect verb acquisition—which may or may not generalize to unstructured interactions (e.g., Naigles et al. 2011; Shulman and Guberman 2007). Although experimental studies are valuable for increasing the internal

validity of findings, translating findings from studies using contrived input conditions to naturally occurring interactions or naturalistic intervention sessions is sometimes challenging.

Second, our study suggests past demonstrations that individual differences in caregiver verb input is associated with typically developing toddlers' verb vocabularies generalize to children with ASD. With the exception of the Crandall and colleagues' (in press) study, the studies examining the association between caregiver verb input within less contrived interactions and later child expressive verb vocabulary were conducted on children with TD, not children with ASD (e.g., Hadley et al. 2016).

Third, the current study expands on findings that follow-in verb input predicted later child expressive verb vocabulary in children with ASD by showing the importance of the engagement state in which the follow-in verb input occurs. Crandall and colleagues (in press) found that follow-in verb input in unstructured interactions with children with ASD predicted later child expressive verb vocabulary, but did not examine the engagement states in which follow-in verb input occurred.

Fourth, the current study elaborates on the findings of one of the only other studies to consider the engagement states in which follow-in utterances occur (Bottema-Beutel et al. 2014) by illustrating the types of interactive contexts necessary for *verb* input and later expressive *verb* acquisition. Bottema-Beutel and colleagues (2014) did not examine specific syntactic categories or types of vocabulary within the follow-in utterances. The current study and Bottema-Beutel et al. (2014) suggest that engagement state needs to be considered in studies examining the relation between caregiver input and language acquisition in children with ASD because engagement state information indicates the types of nonlinguistic contexts that might support children with ASD in learning to talk.

Speculative Explanations for the Findings

Although these findings are correlational, there is sound theory for why a specific association between follow-in verb input in HSJE and later child expressive verb vocabulary occurred. It is possible that children with ASD are best able to connect follow-in verb input with their action referents when they are embedded in an active and sustained interaction centered around actions on objects. For example, a caregiver might say, "Let's build a tower," while stacking a set of blocks. The child then imitates this action (signaling a state of HSJE through imitation). The adult comments, "We're building a tower," (a follow-in utterance with a verb in HSJE) and then takes her turn placing a block on the tower. The child and adult continue to take turns stacking the blocks as the adult provides further follow-in

verb utterances related to building. This reciprocal, active interaction in which the child is an active participant might make the connection between the target verb "build" and its meaning more salient than if the follow-in verb had been presented in another state. For example, these same "build"-focused follow-in utterances could have been used while the child built the tower independently (a state of OE), or while the caregiver influenced the child's play as the child built a tower (e.g., by holding the lower blocks steady), but the child was not collaborating with the caregiver (a state of LSJE). In these states, the child might be less likely to connect the word 'build' with the act of building because he is not reciprocally engaged with the adult. Although such input would still be considered follow-in verbs, its connections to the intended referent actions may be less clear to the child with ASD. Understanding verb meanings provides the semantic basis for accurate verb expression.

Similarly, HSJE engagement formats provide the child with an interactive context to learn how verb constructions are used to interact with another person, not just their understanding of what particular verbs mean. In the above example of HSJE, the caregiver uses a follow-in verb to begin an HSJE turn taking routine ("Let's build a tower"). When follow-in verb input is used in reciprocal exchanges such as these, the child is being given information about how utterances are produced to accomplish social actions (e.g., launching a tower building sequence; Levinson 2013). The child's participation within this action-oriented routine may draw attention to how the verb can be used to interact with the adult. As the child experiences the action-oriented routine, the links between the referent action he is doing or watching, the verb's meaning, and how the verb might be used in meaningful interactional exchanges may be tightened. The reciprocal and social participation in the interaction are not characteristic of LSJE or OE, and may be a critical means for children with ASD to learn how to *use* words that they understand.

Limitations and Future Research

There are several limitations that should be considered when interpreting the results of the present study. First, because of the small sample size and our analyses were post hoc, the findings are considered exploratory (i.e., they may or may not be sample-specific).

Second, like all correlational designs, it is possible that there are unexplored third variable explanations for the observed results. Although we eliminated two possible third variable explanations (i.e., earlier child expressive verb vocabulary and verb input in other states), there are other potential third variable explanations we did not control. For example, we were not able to parse out grammatical informativeness or diversity of follow-in verb input in different

engagement states in our analyses. Because we were constrained by extant, relatively short observations, we did not have sufficient observation time to produce variability across participants for counts of diversity or grammatical informativeness within each engagement state of interest. Future experimental designs might untangle the effects of quantity, diversity, and grammatical informativeness of verb input in HSJE on later child verb vocabulary.

Another possible third variable explanation presents itself because the outcome measure was a parent report. Parents characteristics, such as attentiveness to child behavior, might predispose parents to notice and thus report many verbs and to engage in HSJE + FV.

A third limitation is that the children in our sample were relatively young developmentally, and rarely, if ever, engaged in episodes of coordinated joint engagement. Thus, follow-in verb input in episodes of coordinated joint engagement could not be tested as a potential influence on verb vocabulary.

A related and fourth limitation is that the ASD population is heterogeneous. These findings are not meant to speak to the population at large, but rather a subgroup of children with ASD who are in the early stages of language learning and have a sufficiently large vocabulary (including nouns) that verb acquisition is generally expected. These findings may generalize to those children with ASD who are at developmentally similar stages and language levels to those in our sample. Future research may include samples with greater variability across participants in total vocabulary size and evaluate if there is an interaction between language level and HSJE + FV.

Finally, we did not distinguish between subtypes of verb input and verb vocabulary in our present analyses. For example, it may be that HSJE is a more supportive state to process and learn action verbs than mental state verbs.

Future experimental manipulation of follow-in verbs in HSJE is needed to determine whether the association with child expressive verb vocabulary is causal and to untangle the natural intercorrelation of amount, diversity, and grammatical informativeness of follow-in verb input in HSJE. For example, a between-group experiment in which children capable of HSJE were randomly assigned to business as usual or treatment groups could be conducted. The adult interventionists (staff or caregivers) in the treatment group could be trained to provide follow-in verbs in episodes of HSJE. Expressive verb vocabulary measures using blind examiners and blind coders from this trained group could be compared to those of children in the business as usual control group. Once follow-in verb input in HSJE is established as one of the causal influences on verb vocabulary in children with ASD who can engage in HSJE, then the subgroups of children and subtypes of verbs to which this generalization can be applied should be addressed.

Strengths

There are several strengths of this study that should also be noted. First, the use of a *longitudinal* correlational design provides stronger evidence than a concurrent correlational design that individual variation in early verb input in HSJE is one of the causal influences on child expressive verb vocabulary. This design has evidence for both association and temporal precedence, two of the three assumptions for causal relations (Beakley and Ludlow 1992). Second, the use of a generalized linear model with maximum likelihood estimates and robust standard errors protected against possible undue influence exerted by outlier data points in our small sample. Third, our statistical control of early child expressive verb vocabulary and follow-in verb input in other states eliminated two alternative explanations for the observed relation. Finally, the additional finding that non-verb follow-in utterances in HSJE did not predict child expressive verb vocabulary lent further specificity to the observed findings and improved the falsifiability of the motivating theory for the study.

Implications for Practice

This study supports the strategies used in many naturalistic developmental behavioral interventions (NDBIs; Schreibman et al. 2015). These interventions are in contrast to adult-directed teaching strategies, such as those used in some discrete-trial behaviorist programs. The evidence provided by this study suggests that mere pairing of a targeted verb with its referent may be insufficient for children like those in our sample. For children with ASD in the early stages of verb learning, input may be most processable when situated within longer, higher-quality episodes of sustained engagement (HSJE). Such episodes match the interactional and language processing needs of children with ASD. Additionally, many NDBIs focus on supporting joint engagement generally rather than HSJE specifically. Our findings suggest HSJE has special utility for promoting verb vocabulary development.

Clinicians and caregivers might (a) focus on supporting child behaviors that can promote episodes of HSJE specifically and (b) use follow-in verbs in episodes of HSJE. For example, through use of common NDBI strategies such as repetition, modeling, and direct prompting and reinforcement of turn-taking, adults can teach children with ASD to participate actively in object-oriented activities, which often are examples of HSJE. Another part of the adults' participation in HSJE activities is to actively and reciprocally guide, expand on, and comment on the actions that they and their children are using or could be using.

However, it is crucial to remember that HSJE is a co-constructed engagement state. Although caregivers can

influence the likelihood that their children enter into or continue engaging in HSJE (Bottema-Beutel et al. 2018), the child must still be able to reciprocally engage with the caregiver. Future research is needed to further determine how we might help children grow in ways that enable episodes of HSJE.

One issue worth noting is the use of follow-in verbs in *directives* versus *comments* in HSJE. Both follow-in directives and comments were included in our coding of follow-in verb input in HSJE. Although some clinicians and caregivers may be hesitant to use directives (i.e., guiding play) in the process of teaching these behaviors, recent findings suggest *follow-in* directives are likely to be followed by episodes of HSJE in young children with ASD (Bottema-Beutel et al. 2018) and are positively associated with expressive language in children with ASD (McDuffie and Yoder 2010).

One important nuance the current study did not address, but that needs attention in light of clinical implications, surrounds the importance of grammatical informativeness and diversity. Previous work has detected a positive relation between grammatical informativeness and diversity of verb input in follow-in utterances and later expressive verb vocabulary in children with ASD (Crandall et al., in press). When children with ASD can use morphosyntactic cues to learn verbs, it is important that input be provided in grammatically correct phrases or sentences (Venker et al. 2015). There is little justification for using bare verb stems, even in follow-in utterances in HSJE. Similarly, there is reason to expect that exposing children to many different verbs is more facilitative than using the same few verbs repeatedly (Hsu et al. 2017). Indeed, in at least one sample of children with TD, diversity was a stronger predictor of expressive verb vocabulary than quantity of verb input (Hsu et al. 2017).

Summary

This study is perhaps the first to demonstrate that follow-in verb input in HSJE states is specifically related to later expressive verb vocabulary in at least some children with ASD. Verbs are crucial for building grammatical utterances and developing generative language. Unfortunately, verbs are a commonly understudied aspect of language development for children with ASD. As an early step in a programmatic line of research, this correlational study is designed to identify how verb input might be best provided to support the verb acquisition of children with ASD. It is our hope that this study will spur on further research in methods of facilitating verb vocabulary in children with ASD.

Author Contribution MC conceived of the study, participated in the design, video coding, and data analysis, and drafted the manuscript; KB participated in the video coding, helped conceptualize the study,

and edited the manuscript; JM participated in the video coding, data analysis, and edited the manuscript; LW recruited participants, collected data, participated in the interpretation of the results, and was a site PI on the grant; PY conceived of the study, recruited participants, collected data, guided data analysis, edited the manuscript, and was the primary PI on the grant. All authors read and approved the final manuscript.

Funding This research was funded by National Institute for Deafness and other Communication Disorders (NIDCD R01 DC006893) and supported by the National Institute for Child Health and Disorders (NICHD; P30HD15052; P30HD03110), a US Department of Education Preparing the Next Generation of Scholars in Severe Disabilities grant (H325D140077), and a US Department of Education Preparation of Leadership Personnel grant (H325D140087). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the US Department of Education.

Compliance with Ethical Standards

Conflicts of interest The authors declare they have no conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. (2009). Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism and Developmental Disorders*, *39*, 84–96. <https://doi.org/10.1007/s10803-008-0601-7>.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: American Psychiatric Association.
- Beakley, B., & Ludlow, P. (1992). *The philosophy of the mind: Classical problems/contemporary issues*. Cambridge, MA: The MIT Press.
- Bottema-Beutel, K., Lloyd, B., Watson, L., & Yoder, P. (2018). Bidirectional influences of caregiver utterances and support joint engagement in children with and without autism spectrum disorder. *Autism Research*, *11*, 755–765. <https://doi.org/10.1002/aur.1928>.
- Bottema-Beutel, K., Yoder, P. J., Hochman, J. M., & Watson, L. R. (2014). The role of supported joint engagement and parent utterances in language and social communication development in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*, 2162–2174. <https://doi.org/10.1007/s10803-014-2092-z>.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S., Pascual, L., et al. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, *75*, 1115–1139. <https://doi.org/10.1111/j.1467-8624.2004.00729.x>.

- Brandone, A. C., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development, 78*, 1322–1342. <https://doi.org/10.1111/j.1467-8624.2007.01068.x>.
- Crandall, M., McDaniel, J., Watson, L. R., & Yoder, P. J. (in press). An exploratory analysis of the relationship between parent verb input and verb expressive vocabulary in children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*.
- Dawson, G., Toth, K., Abbot, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*, 271–283. <https://doi.org/10.1037/0012-1649.40.2.271>.
- Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders, 37*, 1007–1023. <https://doi.org/10.1007/s10803-006-0239-2>.
- Erceg-Hurn, D. M., & Mirosevic, V. M. (2008). Modern robust statistical methods: an easy way to maximize the accuracy and power of your research. *American Psychologist, 63*, 591. <https://doi.org/10.1037/0003-066X.63.7.591>.
- Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, J., & Bates, E. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Paul H. Brookes Publishing Inc.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*, 23–64. https://doi.org/10.1207/s15473341l1d0101_4.
- Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Introduction: Progress on the verb learning front. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 3–28). New York: Oxford University Press.
- Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12*, 397–403. <https://doi.org/10.1016/j.tics.2008.07.003>.
- Hadley, P. A., Rispoli, M., & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. *Language, Speech, and Hearing Services in Schools, 47*, 44–58. https://doi.org/10.1044/2015_LSHSS-15-0018.
- Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*, 212–229. <https://doi.org/10.1111/j.1469-7610.2004.00215.x>.
- Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language, 44*, 63–86. <https://doi.org/10.1017/S0305000915000690>.
- Johnson, V. E., & de Villiers, J. G. (2009). Syntactic frames in fast mapping verbs: Effect of age, dialect, and clinical status. *Journal of Speech, Language, and Hearing Research, 52*, 610–622. [https://doi.org/10.1044/1092-4388\(2008/07-0135\)](https://doi.org/10.1044/1092-4388(2008/07-0135)).
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155–163.
- Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103–130). Malden, MA: Wiley-Blackwell.
- Lidz, J. (2006). Verb learning as a probe into children's grammars. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 429–449). New York: Oxford University Press.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule* (2nd ed.). Torrance, CA: Western Psychological Services.
- Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364–391). New York: Oxford University Press.
- McDuffie, A. S., & Yoder, P. (2010). Types of parent verbal responsiveness that predict language in young children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 53*, 1026–1039. [https://doi.org/10.1044/1092-4388\(2009/09-0023\)](https://doi.org/10.1044/1092-4388(2009/09-0023)).
- McGregor, K. K., Licandro, U., Arenas, R., Eden, N., Stiles, D., Bean, A., et al. (2013). Why words are hard for adults with developmental language impairments. *Journal of Speech, Language, and Hearing Research, 58*, 1845–1856. [https://doi.org/10.1044/1092-4388\(2013/12-0233\)](https://doi.org/10.1044/1092-4388(2013/12-0233)).
- Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Los Angeles: Western Psychological Services.
- Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Research, 4*, 422–437. <https://doi.org/10.1002/aur.223>.
- Plante, E., Bahl, M., Vance, R., & Gerken, L. (2011). Beyond phonotactic frequency: Presentation frequency effects word productions in specific language impairment. *Journal of Communication Disorders, 44*, 91–102.
- Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., et al. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*, 2411–2428. <https://doi.org/10.1007/s10803-015-2407-8>.
- Shulman, C., & Guberman, A. (2007). Acquisition of verb meaning through syntactic cues: A comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD). *Journal of Child Language, 34*, 411–423. <https://doi.org/10.1017/S0305000906007963>.
- Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., et al. (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*, 643–652. [https://doi.org/10.1044/1092-4388\(2009/08-0136\)](https://doi.org/10.1044/1092-4388(2009/08-0136)).
- Tapp, J. T. (2003). ProCoderDV [Computer software and manual]. http://procoder.vueinnovations.com/sites/default/files/public/d10/ProCoderDVManual_0.pdf.
- Tapp, J. T. (2013). *ProCoder merger [computer software]*. Nashville, TN: Vanderbilt Kennedy Center.
- Tapp, J., & Walden, T. A. (1993). PROCODER: A professional tape control, coding, and analysis system for behavioral research using videotape. *Behavior Research Methods, Instruments and Computers, 25*, 53–56. <https://doi.org/10.3758/BF03204449>.
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*, 311–333. <https://doi.org/10.1017/S0305000900011430>.
- Venker, C. E., Bolt, D. M., Meyer, A., Sindberg, H., Ellis Weismer, S., & Tager-Flusberg, H. (2015). Parent telegraphic speech use and spoken language in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research, 58*, 1733–1746. https://doi.org/10.1044/2015_JSLHR-L-14-0291.
- Yoder, P., & Kaiser, A. P. (1989). Alternative explanations for the relationship between maternal verbal interaction style and child language development. *Journal of Child Language, 16*, 141–160. <https://doi.org/10.1017/S0305000900013489>.
- 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.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.