Research Article

Acquisition of Verb Meaning From Syntactic Distribution in Preschoolers With Autism Spectrum Disorder

Sabrina Horvath,[a] Elizabeth McDermott,[a]
Kathleen Reilly,[b] and Sudha Arunachalam[a]

Purpose: Our goal was to investigate whether preschool children with autism spectrum disorder (ASD) can begin to learn new word meanings by attending to the linguistic contexts in which they occur, even in the absence of visual or social context. We focused on verbs because of their importance for subsequent language development.
Method: Thirty-two children with ASD, ages 2;1–4;5 (years; months), participated in a verb-learning task. In a between-subjects design, they were randomly assigned to hear novel verbs in either transitive or intransitive syntactic frames while watching an unrelated silent animation or playing quietly with a toy. In an eye-tracking test, they viewed two video scenes, one depicting a causative event (e.g., boy spinning girl) and the other depicting synchronous events (e.g., boy

and girl waving). They were prompted to find the referents of the novel verbs, and their eye gaze was measured.
Results: Like typically developing children in prior work, children with ASD who had heard the verbs in transitive syntactic frames preferred to look to the causative scene as compared to children who had heard intransitive frames.
Conclusions: This finding replicates and extends prior work on verb learning in children with ASD by demonstrating that they can attend to a novel verb's syntactic distribution absent relevant visual or social context, and they can use this information to assign the novel verb an appropriate meaning. We discuss points for future research, including examining individual differences that may impact success and contrasting social and nonsocial word-learning tasks directly.

Many children with autism spectrum disorder (ASD) have impairments in language ability that affect lexical knowledge, including receptive vocabulary (e.g., Charman, Drew, Baird, & Baird, 2003; Ellis Weismer, Lord, & Esler, 2010; Kover, McDuffie, Hagerman, & Abbeduto, 2013; Luyster, Kadlec, Carter, & Tager-Flusberg, 2008; Miniscalco, Fränberg, Schachinger-Lorentzon, & Gillberg, 2012; Mitchell et al., 2006). Receptive vocabulary is a prerequisite for other aspects of language learning, impacting acquisition of grammar (E. Bates & Goodman, 1999; Marchman & Fernald, 2008) and literacy (Scarborough, 2001). Furthermore, receptive vocabulary in toddlerhood is a predictor of long-term outcomes for children

with ASD (e.g., Hudry et al., 2014; Sigman et al., 1999; Venter, Lord, & Schopler, 1992). Therefore, in the current study, we focus on mechanisms underlying the ability to acquire new receptive vocabulary in ASD.

It is likely that the difficulty children with ASD have in acquiring new words is principally related to word meaning rather than word form (e.g., Naigles & Tek, 2017; Norbury, Griffiths, & Nation, 2010). Typically developing children can use a speaker's social communicative intent in order to infer word meaning, but social cognition is a known area of difficulty in ASD (e.g., Baron-Cohen, Baldwin, & Crowson, 1997; Charman, 2003; Warreyn, Roeyers, Wetswinkel, & de Groote, 2007). For example, children with ASD, as compared to language-matched typically developing children, have difficulty in using cues such as the speaker's direction of gaze or pointing to infer the meaning of a novel word (e.g., Baron-Cohen et al., 1997; Norbury et al., 2010; Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff, & Tager-Flusberg, 2007; Preissler & Carey, 2005). However, children with ASD do have relative strengths in acquiring aspects of linguistic form (e.g., Norbury et al., 2010), which we suggest may, in some cases, help them acquire meaning, particularly if social communicative

[a]Boston University, MA
[b]Tufts Medical Center, Floating Hospital for Children, Boston, MA
Correspondence to Sabrina Horvath: shorvath@bu.edu

demands are minimized. We explore this possibility in the current study.

One demonstrated ability in children with ASD is the use of statistical or distributional learning—that is, learning by extracting patterns or regularities from the input. (Note that the terms statistical learning and distributional learning are often used interchangeably in literature. We hereafter use the term distributional learning for simplification.) Children with ASD have been shown to be able to use distributional learning to identify word forms from a speech stream (although they may require more exposure than typically developing children; see Kover, 2018, for discussion of the limited research on children with both ASD and intellectual disability; e.g., Eigsti & Mayo, 2011; Foti, De Crescenzo, Vivanti, Menghini, & Vicari, 2015; Mayo & Eigsti, 2012; Obeid, Brooks, Powers, Gillespie-Lynch, & Lum, 2016). In these studies, children with ASD were presented with a stream of speech in which some syllables co-occurred more often than others, and they successfully identified the ones with high co-occurrence as coherent units or words. However, extracting a linguistic form is but one part of word learning: In order to learn a word, children must pair forms they identify with meaning. Graf Estes, Evans, Alibali, and Saffran (2007) found that typically developing 17-month-olds could map the "words" they had segmented from the speech stream to novel objects (see also Lany & Saffran, 2010). More recently, Haebig, Saffran, and Ellis Weismer (2017) found that school-age children with ASD, too, can map newly segmented word forms to meaning.

A common mechanism believed to be at play for verb learning in particular—syntactic bootstrapping—is at base a kind of distributional learning, coupled with a link from form to meaning (Gleitman, 1990). We are particularly interested in children's verb acquisition abilities because verb vocabulary is a stronger predictor of later outcomes than the better studied noun vocabulary (e.g., Hadley, Rispoli, & Hsu, 2016). In syntactic bootstrapping, learners extract distributional information about a verb's linguistic environment—that is, the distribution of elements within the sentence with respect to each other, such as whether a word is followed by a noun phrase (e.g., John ate cookies) or a prepositional phrase (e.g., John ran to the store). This information constrains the learner's hypothesis space about the verb's meaning; a verb followed by a noun phrase in object position is more likely to label an event involving two event participants, in which an agent acts on a patient, than a verb followed by a prepositional phrase, which is likely to label an event involving one event participant acting alone. This kind of learning differs from the tasks previously described in that the information extracted from these patterns is not about a word's phonological form but rather its grammatical category and syntactic distribution. In order to make use of this information to identify the verb's meaning, children must have knowledge about syntax–semantics links, such as relationships between syntactic positions (e.g., subject, object) and semantic roles (e.g., agent, patient).

By 1.5–2 years of age, typically developing children use syntactic bootstrapping to infer a novel verb's broad category of meaning (Arunachalam, Escovar, Hansen, & Waxman, 2013; Fisher, 1996, 2002; Messenger, Yuan, & Fisher, 2015; Naigles, 1990; Naigles & Kako, 1993; Yuan & Fisher, 2009). In Naigles's (1990) classic study, children were presented with a scene of two actors simultaneously engaged in both a causative event (e.g., duck bends a bunny at the waist) and synchronous, noncausative events (e.g., duck and bunny both flex their arms). Concurrently, they heard a novel verb in an informative syntactic frame (e.g., "Look! The duck is gorping the bunny," Naigles, 1990, p. 363). At test, they saw the causative and synchronous events pulled apart, side by side, and were asked to find "gorping." Children succeeded, preferring the causative scene if they had heard the verb used transitively as compared to if they had heard it used intransitively.

Children with ASD, too, can engage in syntactic bootstrapping (Naigles, Kelty, Jaffery, & Fein, 2011; Shulman & Guberman, 2007). Like the typically developing 2-year-olds in Naigles (1990), 5-year-old Hebrew-acquiring children with ASD (Shulman & Guberman, 2007) and 3-year-old English-acquiring children with ASD (Naigles et al., 2011) mapped novel verbs occurring in transitive frames to causative events, indicating that they used the verb's syntactic frame to infer that the verb labeled the event in which an agent acted on a patient.

In the task used in these two studies, like the original Naigles (1990) study, children were simultaneously presented with both linguistic information (form) and candidate event referent (meaning), making this task different from ones like Graf Estes et al. (2007), in which children had to first extract a form using distributional information alone, maintain this linguistic representation, and subsequently map it to meaning. This kind of dissociation between distributional information and referential meaning is by now also well studied in the context of syntactic bootstrapping in typically developing children. Yuan and Fisher (2009) developed a clever paradigm in which children heard novel verbs in informative syntactic contexts (e.g., transitive frames) in the context of a conversation between two actors. During this period of familiarization to the linguistic context, children did not have access to a candidate referent event; they simply watched the actors converse. Only afterward, at test, did children see causative and noncausative events and were asked to identify the novel verb's referent. Here, too, children succeeded.

In a variant on this paradigm, which we adapt for the current study, Arunachalam (2013) asked if typically developing children could succeed with even fewer social communicative cues than were present in Yuan and Fisher (2009). Instead of being presented in a conversation, the linguistic information was presented in a situation similar to some distributional learning studies of word segmentation (e.g., Graf Estes et al., 2007; Lany & Saffran, 2010)—as a stream of sentences that played as ambient noise while the child played with toys or watched an unrelated silent animation. There was no indication that the sentences were being spoken to the child or to anyone else, and the sentences were presented in a list, as a disconnected series of sentences

rather than as part of a conversation or narrative. Therefore, there was not a clear indication that the novel verbs were relevant to the child or needed to be attended to or remembered. Nevertheless, typically developing children ages 2;1 to 2;5 (years;months) learned the novel verbs. At test, those who had heard transitive frames preferred the causative event as compared to those who had heard intransitive frames.

For children with ASD, this kind of ambient presentation of linguistic information might be particularly helpful, at least for helping them form an initial representation for the new word. After all, given relative strengths with word forms over word meanings (e.g., Norbury et al., 2010), extracting distributional information about the novel word form may be the easy part of word learning for them, and it is possible that the opportunity to do so without simultaneously having to attend to visual or social information might enhance their learning, as some scholars have argued that integrating multiple kinds of information is particularly challenging for children with ASD (e.g., Foxe et al., 2013; Happé & Frith, 2006; Stevenson et al., 2014). On the other hand, we might predict that children with ASD would process the phonological information as they have done in prior distributional learning studies (e.g., Eigsti & Mayo, 2011; Foti et al., 2015; Mayo & Eigsti, 2012; Obeid et al., 2016) but fail to store the verb's syntactic distribution or make the leap to imbuing it with meaning. After all, meaning has a communicative importance that form alone does not. Because the learning situation is not overtly communicative, a new lexical representation may simply not be "worth" forming; children with ASD may instead require a combination of linguistic, visual, and social cues to prompt a transition from form to meaning.

In the current study, we adapted Arunachalam (2013) paradigm to test whether preschool children with ASD (mean age = 3;3) can infer a verb's meaning after hearing it in a stream of transitive or intransitive sentences. We expand on prior syntactic bootstrapping work in ASD in several ways. First, unlike prior work with typically developing children, Naigles et al. (2011) only tested children with ASD in a transitive syntactic condition (e.g., "The boy is mooping the girl"). We instead follow on numerous prior studies with typically developing children (beginning with Naigles, 1990), which used a between-subjects design and randomly assigned children to either a transitive condition or an intransitive condition (e.g., "The boy and the girl are mooping"). Because the intransitive sentence can refer to either a causative event or synchronous events (Arunachalam, Syrett, & Chen, 2016; Naigles & Kako, 1993), the intransitive condition serves as a control against which we compare performance in the transitive condition.[1] Shulman and

Guberman (2007) included both transitive and intransitive conditions, but their study involved children who were quite a bit older (mean age = 5;7), and furthermore, they were acquiring Hebrew, whose syntactic properties differ from English in important ways, making comparison difficult. Shulman and Guberman also used pointing as a measure of learning rather than the eye gaze measures that are more typical of syntactic bootstrapping studies both with typically developing children and children with ASD. We believe that eye-tracking methods are particularly well suited to young children with ASD, as they do not require children to point, vocalize, or carry out actions and therefore may reveal receptive abilities that are otherwise masked (e.g., Naigles & Fein, 2016).

Finally, following Arunachalam (2013) work with typically developing children, we presented the linguistic familiarization in as nonsocial a way as possible, as a stream of unrelated sentences containing the novel verb, in the absence of relevant visual information. This allowed us to examine the following issue: Can children with ASD use representations of linguistic form built from this input as a basis for acquisition of verb meaning? If not, this would suggest that they require an overt, ostensive labeling situation with both visual and communicative context in order to transition from form to meaning.

Method

Participants

Thirty-four children (4 girls, 30 boys) recruited from the greater Boston area were included in the final sample, with an average age of 3;3 (ranging 2;1–4;5, SD = 7 months). Per parent report, all were monolingual English language learners (hearing no more than 20% of another language) with no history of hearing loss. All parents reported that their child had a diagnosis of ASD, autism, or pervasive developmental disorder–not otherwise specified. Given the young age of the children, we did not distinguish among these diagnoses, but we confirmed diagnosis in the lab for all but two children using the Autism Diagnostic Observation Schedule–Second Edition (ADOS-2; Lord et al., 2012) or ADOS–Toddler Module (Luyster et al., 2009), depending on age, either as part of this study's protocol or within 12 months as part of a related project. ADOS-2 testing for 23 of the children was conducted by the third author, a clinician who has received ADOS-2 training and provided a confirmatory diagnosis in conjunction with her expert clinical judgment (most of these children were referred to us by her and received their initial diagnosis from her interdisciplinary team). Testing for nine of the children was conducted by a research-reliable examiner. The remaining two children had a parent-reported diagnosis of ASD but did not participate in confirmatory testing due to scheduling difficulties. These two children's scores on the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) were both above the cutoff of 15 (21 and 23), indicating that they met criterion; however, the SCQ was designed

---

[1]Although we generally think it is ideal to use within-subject manipulations with clinical populations to permit detailed analyses of individual differences, we adopt the standard between-subjects design for this paradigm because of concerns that exposure to one condition would affect performance in the other. See our discussion of the potential impact of this design choice in Arunachalam, Syrett, and Chen (2016).

for slightly older children, and although sensitivity is high for children in our age range, specificity is relatively lower (Allen, Silove, Williams, & Hutchins, 2007).

Most of the children (*n* = 25) were enrolled in applied behavior analysis (ABA), speech therapy, and/or early intervention at the time of participation; 12 were additionally enrolled in physical therapy and/or occupational therapy. Eight families did not provide information regarding their child's therapy history. The children were predominantly White (*n* = 24); the remaining children were African American (*n* = 4), Asian (*n* = 1), or multiracial (*n* = 5); in addition, one child was identified as Hispanic. The majority of mothers (*n* = 20) had a college degree or more advanced degree; of the remaining mothers, four had a high school diploma, six had an associate degree, two had some college, and two families did not provide information regarding maternal education.

Parents completed the MacArthur–Bates Communicative Development Inventories–Short Forms, Level 2 (CDI 2; designed for typically developing children ages 1;4–2;6) and Level 3 (CDI 3; designed for typically developing children ages 2;6–3;1; Fenson et al., 2007). The CDI 2 and CDI 3 checklists provided information about children's expressive vocabularies. Four children in the final sample had incomplete CDI data (one missing CDI 2, two missing CDI 3, and one missing both). On average, children were reported by their parents to produce 50 of the 100 words on the CDI 2 (ranging 0–100, *SD* = 32) and 26 of the 100 words on the CDI 3 (ranging 0–91, *SD* = 28). We also administered the Visual Reception, Receptive Language, and Expressive Language subscales of the Mullen Scales of Early Learning (MSEL; Mullen, 1995). Three children did not participate in any MSEL testing, and a further two only completed two of the three subtests. Of those who participated, 69% had *t* scores more than 1 *SD* below the mean for their chronological age on the Receptive Language subscale, and 76% had *t* scores more than 1 *SD* below the mean for their chronological age on the Expressive Language subscale. This indicates that the majority of participants in our study were delayed in their language development. The mean raw score on the Receptive Language subscale was 26.9 (range 11–40, *SD* = 8), corresponding to an age equivalent of 2;4. Mean raw score on the Expressive Language subscale was 24.0 (range 10–40, *SD* = 8), corresponding to an age equivalent of 2;2. Mean raw score on the Visual Reception subscale was 30.3 (range 12–45, *SD* = 7), corresponding to an age equivalent of 2;5. Taken together, these assessments indicate that the participants in our sample have a language level comparable to a 2-year-old child, consistent with many prior studies of syntactic bootstrapping with typically developing toddlers (e.g., Naigles, 1990, who included participants with a mean age of 2;1). Participants' language levels are also comparable to or slightly below those of the typically developing children tested in the same paradigm in Arunachalam (2013), who had an average chronological age of 2;3 and were reported to produce an average of 75 words on the CDI 2.

Participants were randomly assigned to one of two conditions in a between-subjects design: transitive (*n* = 17) or intransitive (*n* = 17). This between-subjects design is consistent with Arunachalam (2013), as well as prior syntactic bootstrapping studies with typically developing children (e.g., Naigles, 1990; Naigles & Kako, 1993). The two conditions were identical except for the syntactic frame in which novel verbs were presented during the familiarization phase (see below). There were no differences between the conditions with respect to children's average age, the proportion of the group that was male, or in average SCQ raw scores, average CDI 2 raw scores, or average CDI 3 raw scores. In addition, there were no group differences in the average raw score on any of the three MSEL subscales (Visual Reception, Receptive Language, or Expressive Language) or in the proportion of the group with a *t* score of more than 1 *SD* below the mean for any of the subscales (see Table 1).

An additional four participants were excluded from analysis, two because they chose not to watch any of the study videos and the other two (one assigned to the transitive and the other to the intransitive condition) because of high levels of track loss on every trial (see below).

Apparatus and Materials

Apparatus

Videos were displayed on a Tobii T60 XL 24-in. corneal reflection eye-tracking monitor, which samples gaze every 17 ms. Children were seated either in a car seat approximately 20 in. from the monitor or in their parent's lap. If the latter, the parent wore a blindfold.

Stimuli

Each trial consisted of three phases: familiarization, test, and generalization.

Visual stimuli. Visual stimuli during familiarization consisted of animated images of colored shapes and lines that were unrelated to the linguistic content of the experiment.

The test phase and generalization phase each featured two scenes: One scene depicted a causative event (e.g., one actor spinning the other by her shoulders), and the other depicted synchronous (noncausative) events (e.g., two actors waving hands in a clockwise motion). Two of the four trials featured events with two human actors; the other two trials featured events with one actor and one object. Actors (and objects) were the same in both the causative and synchronous scenes of a single trial but changed between the test phase and the generalization phase of each trial. There was no repetition of actors or objects across trials (see Table 2).

Auditory stimuli. Auditory stimuli for this study were produced by a female native American English speaker in a sound-attenuated booth. For the familiarization phase, auditory stimuli consisted of 27 sentences containing a novel verb. Sentences were either consistently transitive (e.g., "Mommy biffed the train") or intransitive (e.g., "Mommy and the train biffed"), corresponding to the child's random

Table 1. Comparison of demographics and SCQ, CDI 2, CDI 3, and MSEL scores in the two syntactic conditions.

Measure	Transitive	Intransitive	Statistic	*p*
Age	*M* = 3;3 (years;months) *SD* = 8 months *n* = 17	*M* = 3;3 (years;months) *SD* = 7 months *n* = 17	*t* = −0.08	*p* = .93, *ns*
Proportion male	*P* = .94 *n* = 17	*P* = .82 *n* = 17	*z* = 1.06	*p* = .28, *ns*
SCQ raw scores	*M* = 16.13 *SD* = 6.1 *n* = 16	*M* = 16.06 *SD* = 6.1 *n* = 16	*t* = 0.03	*p* = .97, *ns*
CDI 2 raw scores	*M* = 52.06 *SD* = 31.0 *n* = 16	*M* = 48.31 *SD* = 34.4 *n* = 16	*t* = 0.32	*p* = .75, *ns*
CDI 3 raw scores	*M* = 25.62 *SD* = 28.3 *n* = 16	*M* = 27.7 *SD* = 15 *n* = 15	*t* = −0.01	*p* = .99, *ns*
MSEL Visual Reception raw scores	*M* = 31.6 *SD* = 9.1 *n* = 16	*M* = 28.7 *SD* = 4.6 *n* = 14	*t* = 1.12	*p* = .27, *ns*
MSEL Visual Reception: Proportion with *t* scores more than 1 *SD* below the *M*	*P* = .63 *n* = 16	*P* = .64 *n* = 14	*z* = 0.10	*p* = .92, *ns*
MSEL Receptive Language raw scores	*M* = 27.3 *SD* = 9.3 *n* = 16	*M* = 26.5 *SD* = 6.1 *n* = 13	*t* = 0.25	*p* = .81, *ns*
MSEL Receptive Language: Proportion with *t* scores more than 1 *SD* below the *M*	*P* = .63 *n* = 16	*P* = .77 *n* = 13	*z* = 0.83	*p* = .40, *ns*
MSEL Expressive Language raw scores	*M* = 24.1 *SD* = 9.2 *n* = 15	*M* = 23.6 *SD* = 8.1 *n* = 14	*t* = 0.15	*p* = .88, *ns*
MSEL Expressive Language: Proportion with *t* scores more than 1 *SD* below the *M*	*P* = .73 *n* = 15	*P* = .79 *n* = 14	*z* = 0.32	*p* = .74, *ns*

Note. SCQ = Social Communication Questionnaire; CDI 2 and CDI 3 = MacArthur–Bates Communicative Development Inventories–Short Forms, Levels 2 and 3, respectively; MSEL = Mullen Scales of Early Learning; *ns* = not statistically significant.

assignment to condition. Sentences varied in the event participants named as well as tense and aspect morphology. The sentences were recorded in adult-directed speech with list-reading prosody, because the goal was to minimize the communicative relevance of the situation for the child as much as possible. There was approximately 1 s between sentences.

In the test and generalization phases, auditory stimuli consisted of attention-grabbing phrases (e.g., "Look!") as well as prompts intended to draw children's attention to the scene depicting the novel verb's referent (e.g., "Where's biffing?").

Procedure

This study was conducted as part of a larger two-visit protocol in our laboratory. At the first visit, families were first brought to a waiting room where children played with experimenters and became comfortable in the lab. Parents provided written consent on behalf of their child and completed paperwork (CDI 2, CDI 3, SCQ). Children and parents were then escorted into the testing room, where children watched the experimental videos on the eye-tracking monitor. The procedure began with a 5-point calibration. Children then saw two warm-up trials and four experimental trials.

Table 2. Description of visual scenes shown during the test phase of each trial.

Novel verb	Test phase	Participants	Causative action	Synchronous action
Biff	1	Two girls	One girl pulling the other by the arms until bent in half at the waist	Both girls bending at the knees, simultaneously
	2	Two (different) girls		
Fez	1	Girl and large green ball	Girl bouncing ball up and down, slowly	Girl, seated, and ball rocking side to side
	2	(Different) girl and large yellow ball		
Lorp	1	Man and white box with lid	Person opening box and peering inside	Person and box simultaneously slipping off of chair and onto floor
	2	Girl and green box with lid		
Moop	1	One boy and one girl	One person turning the other around in a circle	Both people waving their hands in a circular motion
	2	Two girls		

After the study videos, children completed MSEL testing. At the second visit, approximately 1 month later, children saw an unrelated experimental video and participated in ADOS testing to confirm diagnosis.

Warm-Up Trials

Children first participated in two warm-up trials. In each, two animated videos played simultaneously on the screen (e.g., a man eating a cookie, a man dancing). The experimenter prompted children to find one of the two scenes (e.g., "Do you see dancing? Find dancing!"). Children were encouraged to look at the screen but were not required to provide any gestural or verbal responses. The purpose of these two trials was to introduce them to the task of seeing two videos simultaneously while only one was labeled auditorily.
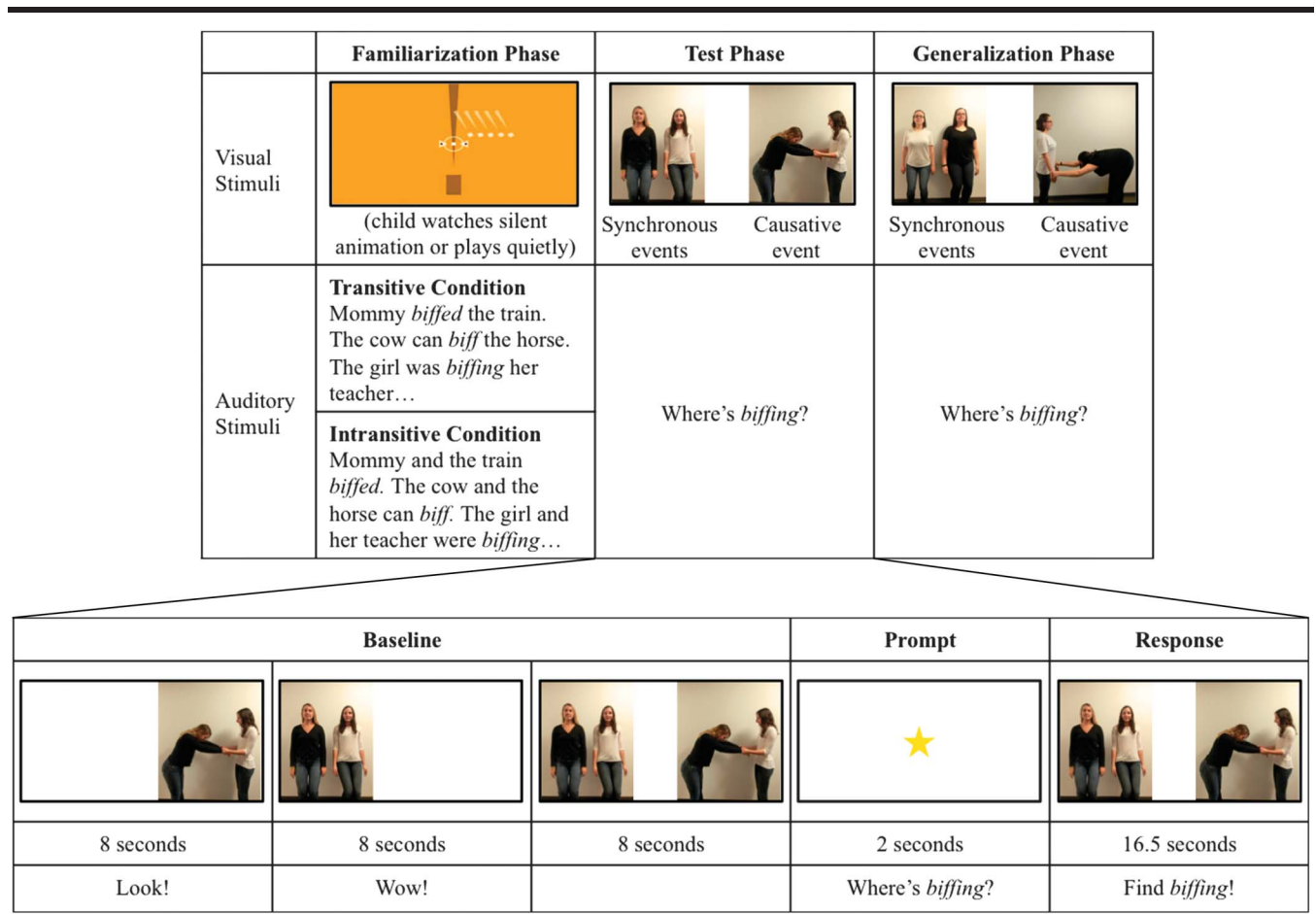
Experimental Trials

Children saw four experimental trials in one of two orders (one the reverse of the other). Each consisted of a familiarization phase, a test phase, and a generalization phase (see Figure 1). The inclusion of the generalization phase is the only design difference between Arunachalam (2013) and this study; we included this to see whether children could not only identify the meaning of the verb in the test phase but also generalize its use to a new visual scene (with different actors and objects) in the generalization phase.

During familiarization (approximately 75 s for each trial), children were introduced to a novel verb in either transitive or intransitive sentences, depending on condition assignment. Simultaneously, an unrelated silent visual animation of shapes and lines played on the eye-tracker monitor; the animation was constant across conditions but differed slightly on each trial. Because no distinguishable events or event participants were displayed, the animation provided no cues to support verb learning; children could only learn from the auditory stimuli. Children were not required to look at the screen during familiarization, and some instead played quietly with small toys.

Next, in the test phase, which was identical across conditions, children viewed two events side by side. The test phase consisted of baseline, prompt, and response subphases. During baseline (25 s), children viewed two candidate referent events for the novel verb; each event was depicted in a separate scene, one on each side of the monitor. Scenes were presented first individually (8 s each

Figure 1. Schematic depiction of stimuli for one trial.

	Familiarization Phase	Test Phase	Generalization Phase
Visual Stimuli	(child watches silent animation or plays quietly)	Synchronous events / Causative event	Synchronous events / Causative event
Auditory Stimuli	**Transitive Condition** Mommy *biffed* the train. The cow can *biff* the horse. The girl was *biffing* her teacher… **Intransitive Condition** Mommy and the train *biffed.* The cow and the horse can *biff.* The girl and her teacher were *biffing*…	Where's *biffing*?	Where's *biffing*?

Baseline			Prompt	Response
8 seconds	8 seconds	8 seconds	2 seconds	16.5 seconds
Look!	Wow!		Where's *biffing*?	Find *biffing*!

separated by a 0.5-s blank screen) and then simultaneously side by side (8 s), so that children could examine each. The scenes were accompanied by attention-getting language (e.g., "Look!") but no novel words. In the prompt subphase (2 s), the two scenes disappeared, replaced by a fixation star in the middle of the screen. Children were asked to find the referent scene for the novel verb (e.g., "Where's biffing?"). In the response subphase (16.5 s), both scenes again played simultaneously (two iterations of 8 s, separated by a 0.5-s blank screen), and an additional prompt was heard after 4 s had elapsed (e.g., "Find biffing!"). Critically, the prompts heard in the test phase incorporated a neutral syntactic frame that provided no information that could disambiguate between the two scenes; to succeed, children had to rely on the representations for the novel verbs that they had posited during the familiarization phase.

The generalization phase (excluded from analysis, for reasons described below) immediately followed the test phase and was structured identically, except that the test scenes differed slightly (different actors, different backgrounds). The side of presentation (left or right) of the causative and synchronous scenes was held constant for the entirety of the test phase (i.e., from the baseline to response subphase within each trial) but counterbalanced across the test and generalization phases and counterbalanced across the four trials.

Analyses

The critical question to be addressed by our analyses is the following: When children are asked, e.g., "Where's biffing?" after being familiarized to the novel verb, do children who had heard the verb in transitive sentences look more to the causative scene than those who had heard the verb in intransitive sentences? Such a difference between conditions would indicate that children's interpretation of the novel verbs is influenced by the linguistic information they had previously heard.

We excluded trials on which children accumulated 50% or more track loss (e.g., blinks) during the first 6 s of the response subphase. We used this criterion to exclude children who were inattentive after the prompt. Two children failed to contribute any trials given this criterion; they were excluded from the sample altogether. The final sample of 34 children contributed three trials each on average (range 1–4), and 23% of trials were excluded.

We applied this same criterion to the generalization phase; this resulted in exclusion of 53% of trials. Six of our 34 participants contributed no trials to the generalization phase, and the remaining 28 participants contributed an average of two trials each. Following the same benchmark as we did for individual trial exclusion, which required 50% of the data to be useable, we determined that we had insufficient data for analysis of the generalization phase. In reviewing videos of the sessions, it seemed to us that children had simply lost interest in viewing such similar visual scenes for such a long time. Therefore, we excluded the

generalization phase from analysis altogether and focused our attention only on the response subphase of the test phase.

In the response subphase, children are expected to look toward the scene they think best depicts the novel verb after being prompted to do so (e.g., "Where's biffing?"—as heard in the prompt subphase immediately before). Similar verb learning studies have suggested that children will require approximately 1 s from the onset of the response subphase to settle on one scene. Therefore, following Arunachalam (2013), we analyzed gaze during the first 3.5 s of the response subphase, predicting no differences between conditions in the first 1 s, but predicting a difference between conditions to emerge in the 1- to 2.5-s window.

We tallied whether each child looked to the causative scene (1) or elsewhere (0) at each frame (~1/60 s); looks to neither scene and track loss (e.g., blinks) were included.[2] Following a now-standard approach (Barr, 2008), we calculated proportions of looking to the causative scene across 50-ms bins (each bin consisting of three frames), transformed the binned data using an empirical logit function, and entered the transformed data into a weighted mixed-effects regression with maximum likelihood estimation using the lme4 package (Version 1.1-12; D. Bates, Maechler, Bolker, & Walker, 2015) in R (Version 3.3.0; R Development Core Team, 2014). (The empirical logit transform is thought to be the most appropriate for proportion data in this kind of statistical analysis; see Barr, 2008; Jaeger, 2008.) Following Arunachalam (2013), we included random intercepts for participant and trial[3] and fixed effects for time (in seconds), condition (transitive, contrast coded as 0.5, or intransitive, contrast coded as −0.5), time window (0–1 s, contrast coded as −0.5, or 1–2.5 s, contrast coded as 0.5), and a Condition × Time × Time Window interaction. Model comparison was done using the drop1() function with chi-square tests (each parameter is dropped from the model; if the model is better fitting with the parameter than without, we infer that the parameter contributes significantly to the model). Given our prediction that a difference between conditions would emerge in the second time window, the key question was whether model comparison would show a significant

---

[2]Although some studies using similar paradigms (looking while listening, Fernald, McRoberts, & Swingley, 2001; intermodal preferential looking paradigm, Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987) exclude all looks that are not either the target or distractor, we included these looks following research in the Visual World Paradigm tradition (see, e.g., Barr, 2008), in part because we think these looks are informative— particularly for children with ASD who may spend more time in off-task gaze behavior, though they appear not to in our task—and in part because we included a central fixation before the Response subphase, meaning that excluding looks to neither scene would have led to a substantial amount of missing data as children shifted their gaze from the center to one of the scenes.

[3]We used these random effects to match our analyses to Arunachalam's (2013) study with typically developing children. Analyses with a richer random-effects structure including a random slope for time show the same pattern.

contribution of the interaction between condition and time window.

Results

Gaze behavior during the first 3.5 s of the response subphase is depicted in Figure 2 as a proportion of looking to the causative scene with all looks (including track loss and looks to neither scene) included in the denominator of the proportion. Inspection of the figure reveals, as expected, no difference between the two conditions in the first second of the response subphase and a difference in the expected direction emerging at 1 s, with children in the transitive condition preferring the causative scene as compared to those in the intransitive condition. This affirms our decision to use the same time window as that used with typically developing 2-year-olds in Arunachalam (2013). Although children with ASD are generally slower language processors than typically developing children (e.g., Bavin & Baker, 2016; Schuh, Eigsti, & Mirman, 2016), those in our sample are roughly matched in language development to the typically developing children in Arunachalam (2013), aged 2;3, as evidenced by our sample's age-equivalent scores on the MSEL language measures (2;4 for Receptive Language; 2;2 for Expressive Language). It is perhaps not surprising, then, that the participants in our study show a similar processing time course. As predicted, the mixed-effects model specified above yielded significant main effects of condition and time window and a significant interaction between them. The positive value of the parameter estimate for condition reveals that looking to the causative scene was greater in the transitive condition (because transitive was contrast coded as +0.5) than the intransitive condition; the positive value of the parameter estimate for time window reveals that looking to the causative scene was greater in the second time window than the first (because the second time window was contrast coded as +0.5), and the interaction indicates that the difference between conditions was greater in the second time window than the first. There is also a significant effect of time; this simply reflects the fact that looking to the causative scene increases over time regardless of condition (this is expected given that it began at 0 at the beginning of the response window due to the previously presented central star). In addition, there is a significant effect of time by time window; this indicates that children's looking patterns across time are different in the first time window than in the second. This again is expected: During the first time window, children are shifting their eyes from a central fixation star to the dynamic scenes, whereas in the second time window children are looking primarily at the dynamic scenes. Parameter estimates from the model are listed in Table 3.

Given the possibility that children with ASD might be more likely to engage in off-task gaze behavior and these values were included in the analysis, we report on the proportion of track loss as well as looks to neither the causative nor synchronous scene in comparison to the typically

Figure 2. Children's looking to the causative scene over time (out of all data points, including looks to the synchronous scene, track loss data points, and looks to neither scene) from the onset of the response window, separated by condition. Error bars indicate standard error of participant means. The 1- to 2.5-s window in which we expect differences to emerge is outlined by a dashed box.
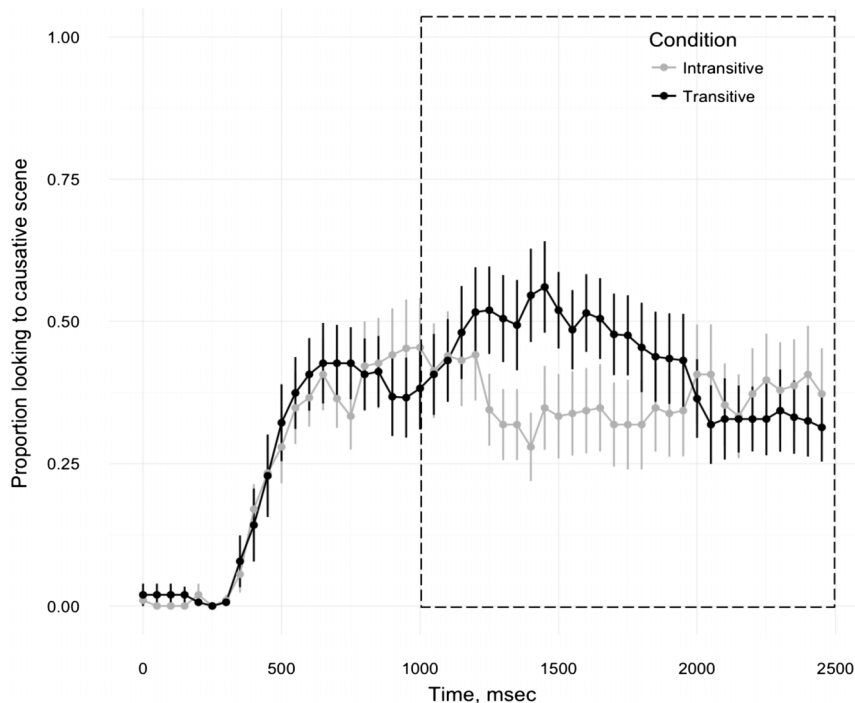
Table 3. Parameter estimates from the best-fitting weighted mixed-effects regression model of the effect of condition and time window on children's preference for the causative scene during the response subphase (empirical logit transformed).

Parameter	Estimate	SD	t	p
Intercept	−1.00	0.12	−8.44	
Time (s)	0.94*	0.075	12.62	< .001
Condition	0.46*	0.22	2.035	.044
Time window	2.30*	0.15	15.77	< .001
Condition × Time window	0.73*	0.29	2.48	.013
Time (s) × Condition	−0.26	0.15	−1.72	.086
Time (s) × Time window	−2.48*	0.15	−16.61	< .001
Time (s) × Condition × Time window	−0.15	0.30	−0.50	.62

Note. Asterisks indicate that the model is better fitting with the parameter than without, as indicated by model comparison.

developing children in Arunachalam (2013). Of the included trials, the proportion of track loss during this first 3.5 s was .10 (*SD* of participant means = 0.11), with no group differences (transitive: *M* = 0.13, *SD* = 0.14; intransitive: *M* = 0.09, *SD* = 0.08; *t* = −0.99, *p* = .33). The proportion of looking to either the causative or the synchronous scene was .70 (*SD* = 0.11) and to neither scene .20 (*SD* = 0.08). Although not reported in that paper, we calculated these percentages for the data in Arunachalam (2013) for comparison: During the first 3.5 s of the response subphase for included trials, the proportion of track loss was .24 (*SD* = 0.18), the proportion of looking to either the causative or synchronous scene was .62 (*SD* = 0.13), and the proportion of looking to neither scene was .13 (*SD* = 0.12).

In our design and in the other studies using this paradigm, the baseline subphase is not a true baseline because children had already been exposed to the novel verb in either transitive or intransitive sentences (see Naigles & Kako, 1993, for discussion). This is why our central comparison is whether there is a difference between conditions in the response subphase as opposed to a change in preference from baseline to response. Nevertheless, we note that preliminary analyses serving as a check revealed no difference between conditions in looking to the causative scene during baseline (β = −.084, *t* = 0.61, *p* = .54), indicating that the difference between conditions in the response subphase is in response to the test query.

Conclusion

The goal of this study was to examine whether children with ASD, like typically developing children, would use the syntactic information available in an ambient auditory stream to not only identify a novel verb's form but also begin to assign it a meaning using syntactic bootstrapping. The children in our study succeeded, demonstrating an ability to use syntactic bootstrapping even though the linguistic information was presented in the absence of any social or visual cues to connect it to an event referent.

Our results build on three prior findings: first, that children with ASD can extract distributional information from an auditory stream (e.g., Eigsti & Mayo, 2011; Foti et al., 2015; Mayo & Eigsti, 2012; Obeid et al., 2016); second, that they can fast-map meaning to a word form segmented by statistical learning (Haebig et al., 2017); and third, that they can use syntactic bootstrapping to learn novel verbs (Naigles et al., 2011; Shulman & Guberman, 2007). Thus, we have demonstrated that (at least some) children with ASD can use syntactic information to establish a representation for a new verb when the auditory stream provides information only about syntactic distribution and when mapping to meaning requires knowledge of syntax–semantics links. Children established representations for the new verbs from their syntactic distribution alone, even without simultaneous access to candidate referents, and added meanings to these representations when visual referents subsequently become available.

This evidence of common learning mechanisms in typically developing children and children with ASD provides additional support to prior reports of syntactic bootstrapping in both populations and expands our understanding of what capacities are affected by the disorder. Children's success in this nonsocial verb-learning task has two possible interpretations. One possibility is that children regularly participate in syntactic bootstrapping as a task of distributional learning, engaging in this process even without a clear communicative context. There are many other kinds of word learning tasks in which typically developing children will only ascribe meaning to a novel word if there is a clear communicative relevance to doing so (e.g., Fennell & Waxman, 2010; Ferguson & Waxman, 2016, 2017; Henderson, Sabbagh, & Woodward, 2013). However, the fact that both typically developing children and children with ASD succeeded in the present task indicates that children can engage in syntactic bootstrapping even if there is no communicative relevance.

Alternatively, communicative relevance may indeed be necessary for vocabulary learning, but by the time that children are advanced enough in their language development that they are able to use syntactic bootstrapping (i.e., their language is equivalent to a typically developing 1.5- to 2-year-old), they recognize the communicative benefit of less social situations like the one we presented. Here, immediate communicative relevance is not necessary to scaffold learning, but communicative relevance more broadly underpins the process and motivates attention to word-learning opportunities.

Irrespective of which of these interpretations is correct, it appears that children with ASD do not need rich social contexts in order to acquire (at least some) word meanings. One intriguing possibility is that they may even perform better when social information is minimized. There is some evidence that minimizing social pressure or social demands supports performance in ASD, for skills such as learning to imitate sounds or interpret a social situation (DeThorne, Johnson, Walder, & Mahurin-Smith, 2009; Pepperberg & Sherman, 2000, 2002; Pierce, Glad, & Schreibman, 1997).

In future work, it will be important to directly compare word learning in nonsocial tasks like ours with learning in more social versions of the task. If children with ASD perform better in less social contexts, there would be strong clinical implications, suggesting that children with ASD might benefit from learning situations in which social demands are minimized (Arunachalam & Luyster, 2016).

Indeed, that children with ASD can use syntactic bootstrapping in nonsocial word learning tasks may ultimately open new avenues for intervention. We offer here some preliminary questions, recognizing that the first, critical step in this process is to directly compare otherwise identical social and nonsocial word learning tasks. Do children with ASD actually perform better in nonsocial contexts or merely just as well? If they perform better, what elements of nonsocial presentation are relevant (e.g., lack of social pressure from an interlocutor, presentation of linguistic cues instead of social cues like pointing or gaze)? How do they fare with outcome measures such as retention and generalization to expressive language given nonsocial presentation contexts? Depending on what, if any, specific elements support learning, retention, and generalization, we may be able to incorporate such elements in intervention. For example, if lack of social pressure to participate in an interaction is key (as happens to be a design element of many experimental word learning tasks as described above; e.g., Arunachalam, 2013; Naigles, 1990; Yuan & Fisher, 2009), it would be useful to pursue these findings in clinical research on group learning contexts in which children can witness learning interactions without necessarily taking part. If other aspects of our nonsocial presentation, such as the absence of a referential context, are useful for helping children to form an initial lexical representation for a new word, we might pursue clinical research in which preliminary exposure to new vocabulary is given via specially designed computer or television media, prior to standard interactive therapy approaches.

Although this preliminary work has focused on an early-emerging linguistic cue as the basis for our distributional learning task, the considerations for social versus nonsocial learning reach far beyond the preschool years. Word learning is not the only language-learning task in which children with ASD may derive benefits from nonsocial presentation of new linguistic material, and nonsocial input may be particularly beneficial for aspects of language for which distributional "form" information can help bootstrap children into meaning (e.g., tense/aspect morphology; Tovar, Fein, & Naigles, 2015; see also Plante & Gómez, 2018, for a discussion of the role of statistical learning in intervention). Future research should also consider these types of language-learning tasks.

Limitations

Given the potential for future research on nonsocial instruction to inform intervention strategies, one limitation of this work is that it identifies group abilities but does not provide any insights into individual performance variation.

Our sample size was too small to permit adequately powered analyses of how children's language or cognitive ability related to task performance. Given the heterogeneity of language abilities in children with ASD, it is likely that optimal contexts for verb learning are not uniform across the population.

Furthermore, although our sample included both preverbal children and above-average language learners, the heterogeneity of our sample should not be misconstrued as evidence that all children across the spectrum are equally likely to benefit from nonsocial linguistic input for syntactic bootstrapping. We agree with the important point made by Naigles et al. (2011) that the syntactic bootstrapping mechanism has many prerequisite abilities and that we would only predict success in children who have these abilities. Children must know some nouns (and arguably, some of the specific nouns used in the contexts being studied; e.g., Sheline & Arunachalam, 2013), they must have both the parsing and grammatical skills to build up a syntactic representation for the sentence in which the unfamiliar verb occurs, and they must have knowledge of syntax–semantics links (e.g., transitive sentences can describe causative events). Naigles et al. (2011) and Shulman and Guberman (2007) demonstrate that there is no reason to believe these abilities will be selectively impaired in ASD (or at least, it seems likely that these abilities will be commensurate with children's receptive language abilities more generally). Therefore, it is likely that only children whose receptive language skills are equal to or more advanced than that of a typically developing 1.5- to 2-year-old will succeed in tasks like ours.

Although we did not have a sufficiently large sample size to look for individual differences, a caveat for future work attempting to do so is that it is often the case that tasks that provide insight into group differences are poor at revealing individual differences (Hedge, Powell, & Sumner, 2017), as tasks (like syntactic bootstrapping) that typically yield relatively large effects of condition have relatively low between-subjects variability; therefore, they necessarily have low reliability for individual differences. Nevertheless, it is likely that nonsocial word learning contexts may not be useful for all children, and identifying such individual differences using tasks that bring out individual variability is an important focus for future work.

Finally, we initially hoped to be able to test children with ASD in a generalization task, in which they had to identify a second instance of the novel verb's referent after the initial test. However, it was clear that the visual scenes did not sufficiently keep their interest through the generalization period. Because generalization is a key aspect of word learning, this will be critical to test in future work with, perhaps, a faster, more engaging test structure.

Acknowledgments

References

Allen, C. W., Silove, N., Williams, K., & Hutchins, P. (2007). Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. *Journal of Autism and Developmental Disorders, 37*(7), 1272–1278. https://doi.org/10.1007/s10803-006-0279-7

Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition, 129,* 569–573. https://doi.org/10.1016/j.cognition.2013.08.021

Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. R. (2013). Out of sight, but not out of mind: 21-Month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes, 28*(4), 417–425. https://doi.org/10.1080/01690965.2011.641744

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

Arunachalam, S., Syrett, K., & Chen, Y. (2016). Lexical disambiguation in verb learning: Evidence from the conjoined-subject intransitive frame in English and Mandarin Chinese. *Frontiers in Psychology, 7,* 138. https://doi.org/10.3389/fpsyg.2016.00138

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*(1), 48–57. https://doi.org/10.1111/j.1467-8624.1997.tb01924.x

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, E., & Goodman, J. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–70). Mahwah, NJ: Erlbaum.

Bavin, E. L., & Baker, E. K. (2016). Sentence processing in young children with ASD. In L. Naigles (Ed.), *Innovative investigations of language in autism spectrum disorder* (pp. 35–48). Washington, DC: American Psychological Association.

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 358*(1430), 315–324. https://doi.org/10.1098/rstb.2002.1199

Charman, 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. https://doi.org/10.1017/S0305000902005482

DeThorne, L. S., Johnson, C. J., Walder, L., & Mahurin-Smith, J. (2009). When "Simon Says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology, 18*(2), 133–145. https://doi.org/10.1044/1058-0360(2008/07-0090)

Eigsti, I. M., & Mayo, J. (2011). Implicit learning in ASD. In D. Fein (Ed.), *The neuropsychology of autism* (pp. 267–279). Oxford, United Kingdom: Oxford University Press.

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. https://doi.org/10.1007/s10803-010-0983-1

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81*(5), 1376–1383. https://doi.org/10.1111/j.1467-8624.2010.01479.x

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur–Bates Communicative Development Inventories*. Baltimore, MD: Brookes.

Ferguson, B., & Waxman, S. R. (2016). What the [beep]? Six-month-olds link novel communicative signals to meaning. *Cognition, 146,* 185–189. https://doi.org/10.1016/j.cognition.2015.09.020

Ferguson, B., & Waxman, S. R. (2017). Linking language and categorization in infancy. *Journal of Child Language, 44*(3), 527–552. https://doi.org/10.1017/S0305000916000568

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in recognizing and understanding words in fluent speech. In J. Weissenborn & B. Höhle (Eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (pp. 97–124). Philadelphia, PA: Benjamins.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology, 31*(1), 41–81. https://doi.org/10.1006/cogp.1996.0012

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55–64. https://doi.org/10.1111/1467-7687.00209

Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D., & Vicari, S. (2015). Implicit learning in individuals with autism spectrum disorders: A meta-analysis. *Psychological Medicine, 45*(5), 897–910. https://doi.org/10.1017/S0033291714001950

Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H. P., Russo, N. N., Blanco, D., . . . Ross, L. A. (2013). Severe multisensory speech integration deficits in high-functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex, 25*(2), 298–312. https://doi.org/10.1093/cercor/bht213

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*(1), 23–45. https://doi.org/10.1017/S030500090001271X

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Hadley, P. A., Rispoli, M., & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. *Language, Speech, and Hearing Services in Schools, 47*(1), 44–58. https://doi.org/10.1044/2015_LSHSS-15-0018

Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry, 58*(11), 1251–1263. https://doi.org/10.1111/jcpp.12734

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods,* 1–21. https://doi.org/10.3758/s13428-017-0935-1

Henderson, A. M., Sabbagh, M. A., & Woodward, A. L. (2013). Preschoolers' selective learning is guided by the principle of relevance. *Cognition, 126*(2), 246–257. https://doi.org/10.1016/j.cognition.2012.10.006

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. https://doi.org/10.1007/s10803-013-1861-4

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Kover, S. T. (2018). Distributional cues to language learning in children with intellectual disabilities. *Language, Speech, and Hearing Services in Schools, 49*(3S), 653–667. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0128

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. https://doi.org/10.1007/s10803-013-1823-x

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science, 21*(2), 284–291. https://doi.org/10.1177/0956797609358570

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule–Second Edition (ADOS-2).* Los Angeles, CA: Western Psychological Services.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., . . . Richler, J. (2009). The Autism Diagnostic Observation Schedule–Toddler Module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(9), 1305–1320. https://doi.org/10.1007/s10803-009-0746-z

Luyster, 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. https://doi.org/10.1007/s10803-007-0510-1

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11*(3), F9–16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Mayo, J., & Eigsti, I. M. (2012). Brief report: A comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. *Journal of Autism and Developmental Disorders, 42*(11), 2476–2485. https://doi.org/10.1007/s10803-012-1493-0

Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development, 11*(4), 356–368. https://doi.org/10.1080/15475441.2014.978331

Miniscalco, C., Fränberg, J., Schachinger-Lorentzon, U., & Gillberg, C. (2012). Meaning what you say? Comprehension and word production skills in young children with autism. *Research in Autism Spectrum Disorders, 6*(1), 204–211. https://doi.org/10.1016/j.rasd.2011.05.001

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 & Behavioral Pediatrics, 27*(2), S69–S78.

Mullen, E. M. (1995). *Mullen Scales of Early Learning.* Circle Pines, MN: AGS.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development, 64*(6), 1665–1687. https://doi.org/10.1111/j.1467-8624.1993.tb04206.x

Naigles, L. R., & Fein, D. (2016). Looking through their eyes: Tracking early language comprehension in ASD. In L. R. Naigles (Ed.), *Innovative investigations of language in autism spectrum disorder* (pp. 49–69). Washington, DC: American Psychological Association.

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*(6), 422–437. https://doi.org/10.1002/aur.223

Naigles, L. R., & Tek, S. (2017). "Form is easy, meaning is hard" revisited: (Re)characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(4), e1438. https://doi.org/10.1002/wcs.1438

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

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology, 7,* 1245. https://doi.org/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. https://doi.org/10.1111/j.1467-8624.2007.01065.x

Pepperberg, I. M., & Sherman, D. (2000). Proposed use of two-part interactive modeling as a means to increase functional skills in children with a variety of disabilities. *Teaching and Learning in Medicine, 12*(4), 213–220. https://doi.org/10.1207/S15328015TLM1204_10

Pepperberg, I. M., & Sherman, D. (2002). A two-trainer modeling system to engender social skills in children with disabilities. *International Journal of Comparative Psychology, 15,* 138–153.

Pierce, K., Glad, K. S., & Schreibman, L. (1997). Social perception in children with autism: An attentional deficit? *Journal of Autism and Developmental Disorders, 27*(3), 265–282. https://doi.org/10.1023/A:1025898314332

Plante, E., & Gómez, R. L. (2018). Learning without trying: The clinical relevance of statistical learning. *Language, Speech, and Hearing Services in Schools, 49*(3S), 710–722. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0131

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*(1), B13–B23. https://doi.org/10.1016/j.cognition.2005.01.008

R Development Core Team. (2014). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Rutter, M., Bailey, A., & Lord, C. (2003). SCQ. *The Social Communication Questionnaire.* Torrance, CA: Western Psychological Services.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice. In F. Fletcher-Campbell, J. Soler, & G. Reid (Eds.), *Approaching difficulties in literacy development: Assessment, pedagogy and programmes* (pp. 23–38). London, United Kingdom: Sage.

Schuh, J. M., Eigsti, I. M., & Mirman, D. (2016). Discourse comprehension in autism spectrum disorder: Effects of working memory load and common ground. *Autism Research, 9*(12), 1340–1352. https://doi.org/10.1002/aur.1632

Sheline, L., & Arunachalam, S. (2013). *When the gostak distims the doshes: Novel verb learning from novel nouns.* Poster presented at the Cognitive Development Society Biennial Meeting, Memphis, TN.

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*(2), 411–423. https://doi.org/10.1017/S0305000906007963

Sigman, M., Ruskin, E., Arbelle, S., Corona, R., Dissanayake, C., Espinosa, M., . . . Robinson, B. F. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development, 64*(1), 1–114. https://doi.org/10.1111/1540-5834.00001

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., . . . Wallace, M. T. (2014). Evidence for diminished multisensory integration in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(12), 3161–3167. https://doi.org/10.1007/s10803-014-2179-6

Tovar, A. T., Fein, D., & Naigles, L. R. (2015). Grammatical aspect is a strength in the language comprehension of young children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 58*(2), 301–310. https://doi.org/10.1044/2014_JSLHR-L-13-0257

Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry, 33*(3), 489–597. https://doi.org/10.1111/j.1469-7610.1992.tb00887.x

Warreyn, P., Roeyers, H., Wetswinkel, U., & de Groote, I. (2007) Temporal coordination of joint attention behavior in preschoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37,* 501–512. https://doi.org/10.1007/s10803-006-0184-0

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*(5), 619–626. https://doi.org/10.1111/j.1467-9280.2009.02341.x