Author manuscript

Published in final edited form as:

J Autism Dev. Disord. 2021. doi:10.1007/s10803-021-05321-0.

Use of Mutual Exclusivity and its Relationship to Language Ability in Toddlers with Autism Spectrum Disorder

Janine Mathée-Scott[1,2], Caroline Larson[1,7], Courtney Venker[2,4], Ron Pomper[3,6], Jan Edwards[1,5], Jenny Saffran[3], and Susan Ellis Weismer[1]

1 Department of Communication Sciences and Disorders and Waisman Center, University of Wisconsin-Madison, 1500 Highland Ave, Madison, WI 53705, USA

2 Waisman Center, University of Wisconsin-Madison, 1500 Highland Ave, Madison, WI 53705, USA

3 Department of Psychology and Waisman Center, University of Wisconsin-Madison, 1500 Highland Ave., Madison, WI 53705, USA

4 Department of Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Road, East Lansing, MI 48824, USA

5 Department of Hearing and Speech Sciences, Maryland Language Science Center, University of Maryland-College Park, 0121 Taliaferro Hall, Chapel Drive, College Park, MD 20724, USA

6 Boys Town National Research Hospital, 555 North 30th St, Omaha NE 68131, USA

7 Department of Pyschological Sciences, University of Connecticut, 406 Babbidge Road, U-1020, Storrs, CT 06269, USA

Abstract

To efficiently learn new words, children use constraints such as mutual exclusivity (ME) to narrow the search for potential referents. The current study investigated the use of ME in toddlers with autism spectrum disorder (ASD) and neurotypical (NT) peers matched on nonverbal cognition. Thirty-two toddlers with ASD and 26 NT toddlers participated in a looking-while-listening task. Images of novel and familiar objects were presented along with a novel or familiar label. Overall, toddlers with ASD showed less efficient looking toward a novel referent when a novel label was presented compared to NT toddlers, controlling for age and familiar word knowledge. However, toddlers with ASD and higher language ability demonstrated more robust use of ME than those with lower language ability.

Introduction

Researchers have long hypothesized that, in order to narrow the possibilities for mapping labels to objects in their environments, children employ specific strategies and constraints on word meanings. For example, infants and toddlers tend to abide by the whole object principle, interpreting labels as mapping to the object as a whole, rather than its parts (Markman, 1990). Additionally, early language learners assume one-to-one object label mappings, which can be observed by their tendency to map a novel label to a novel object rather than to an object for which they already have a label (Markman & Wachtel, 1988). This tendency, known as "mutual exclusivity" (ME), allows children to begin to efficiently disambiguate their language environment and solve the puzzle of mapping label to meaning (Markman & Wachtel, 1988; Markman, 1990). For example, imagine a naturalistic play scenario in which a parent and young child are interacting with a car and garage toy set. The child is interacting with a toy car, for which she already knows the label, "car." Her father then labels the long, slide-like object within the toy set, for which she does not have a label. The father suggests, "Let's put the car on the *ramp*," performing this action as he narrates it. Under the mutual exclusivity bias, the child would use the physical and social context clues (father moves the car onto the long, slide-like object) and her existing lexical knowledge of the word "car" and makes the assumption that the label "ramp," likely belongs to the long apparatus onto which her father placed the car.

While there has been extensive research on these mechanisms in typical development, it remains unclear whether all children use them in the same way. Children who struggle to develop language at the same pace as their peers may struggle as a result of impaired ability to utilize mechanisms such as ME in order to efficiently process and learn language. Children with autism spectrum disorder (ASD), for example, may have intact or impaired structural language ability, and some will present with concomitant structural language impairment (Anderson et al., 2007; Tager-Flusberg & Kasari, 2013; Arunachalam & Luyster, 2016). This variation in structural language ability may be explained in part by individual differences in underlying language learning mechanisms. Such early lexical constraints have been studied across a wide range of developmental stages of typical development (Bion, et al., 2013; Golinkoff et al., 1994) as well as in older children with ASD (deMarchena et al., 2011; Hartley et al., 2019). However, there has been little work to date focused on the use of ME in young children with ASD. Understanding the early development of ME in children with ASD and their relationship to language ability may provide important insight into early differences in vocabulary and structural language development in this population. The current study investigates whether toddlers with ASD follow a similar trajectory to neurotypical (NT) toddlers in the development of this lexical constraint.

Theoretical accounts of Mutual Exclusivity

There have been multiple proposed theoretical accounts for Mutual Exclusivity, beginning with Markman and Wachtel's seminal (1988) paper, which adopted a lexical constraint, bias-based account. Others have proposed a more domain-general account, based in logical reasoning. Halberda, (2003) theorized that mutual exclusivity reflects a logical reasoning in which children are presented with the problem: "A or B," for which their solution is "if not A, therefore B." Others propose a pragmatic account. This account suggests that children disambiguate referents using social cues, such as eye gaze or gesture, which may override logic or constraint-based cues (Clark, 1990; Tomasello & Akhtar, 1995; Diesendruck & Markson, 2001). The pragmatic account suggests that shared knowledge and speakers' communicative intentions drive the observed phenomenon of mutual exclusivity. More recently, verbal, probabilistic, accounts have been proposed. Some have used computational models to account for learning patterns, including ME (Regier, 2005). McMurray and colleagues (2012), propose a dynamic associative model, by which referent selection and word learning occurs through associative, Hebbian learning. Generally, there is not wide theoretical consensus to date as to whether ME is a domain-specific, linguistic mechanism or a more domain-general, cognitive process.

Mutual Exclusivity in typical language development

The developmental trajectory and individual differences in use of ME within typical development has been extensively investigated across multiple paradigms and age groups (e.g., Markman & Wachtel 1988; Golinkoff et al., 1994; Markman et al., 2003). While research has demonstrated basic use of ME as early as 10 months of age (Mather & Plunkett, 2010), robust use of the ME bias and retention of mapped labels has been demonstrated slightly later, in the second year of typical development (Bion et al., 2013). Changes in the use of ME across the typical development trajectory have also been explored. Research suggests that vocabulary size, but not age in months, is a reliable predictor of ME (Law & Edwards, 2015; Lewis et al., 2019). Moreover, children with larger productive vocabularies, regardless of age, tended to have higher accuracy in their use of ME. This suggests that linguistic experience, particularly productive vocabulary, may be an important factor in the development of ME (Lewis et al., 2019; Grassmann et al., 2015). While the directionality of this relationship remains unclear, these findings highlight the importance of further inquiry into how atypical or delayed language learners, who may have decreased vocabulary sizes, might exhibit a different trajectory of ME use.

Mutual Exclusivity in ASD

In recent years, researchers have explored how different mechanisms, including statistical learning (Van Zeeland et al.,2010; Haebig et al., 2017), attentional allocation (Thorup et al., 2017), and word learning constraints including ME may affect lexical acquisition for children with ASD. School-age children with ASD tend to select a novel referent over a familiar one when presented with a novel label, consistent with ME (Preissler & Carrey, 2005). Further work suggests that this behavior is driven by familiar word knowledge and not simply a preference for novelty. deMarchena and colleagues (2011) found that, in an experiment designed to mitigate novelty bias, school-age children (mean age 8 years) and adolescents (mean age 15 years) with ASD demonstrated intact abilities to utilize ME. Moreover, the authors found that the children in both groups who had larger receptive vocabularies more consistently treated words as mutually exclusive than those with smaller receptive vocabulary sizes (de Marchena et al., 2011). However, it is worth noting that this study only investigated use of ME in "high functioning" children and adolescents with ASD (based on language and social functioning severity scores). deMarchena and colleagues (2011) specifically recruited children with "language abilities that were approximately at chronological age level" according to parent report (p. 100). Given that this study found an important relationship between vocabulary size and use of ME, the authors acknowledge that it remains unclear whether children with higher symptom severity or those with lower language ability will utilize ME (de Marchena et al., 2011). As such, studies of ME in ASD populations to date have largely excluded those children on the spectrum at risk for concomitant language impairment. More recently, Hartley, Bird and Monaghan (2019) found that school-aged children with ASD performed similarly to neurotypical (NT) children matched on receptive vocabulary and nonverbal IQ in their use of mutual exclusivity in a referent selection context, but showed reduced accuracy on measures of delayed retention and generalization.

A study by Bedford and colleagues (2013) examined the ME constraint in toddlers who were deemed at increased genetic risk for ASD (based on having an older sibling with a confirmed diagnosis) and found that these toddlers were able to use the ME constraint to select the correct referent in their task. However, research suggests that only 10-20% of younger siblings of children with confirmed ASD will go on to receive a diagnosis of ASD themselves (Szatmari et al., 2016). Taken together, the literature to date suggests that school-age children with ASD who have average or above average language ability appear to demonstrate use of the ME constraint. Given that there appear to be individual differences within this population in the robustness of the effect, and that this effect is associated with vocabulary attainment (deMarchena et al., 2011), there is a clear need for more information about the use of ME in younger children with confirmed ASD. Further, additional research is needed to increase understanding of lexical acquisition mechanisms, such as ME, in younger children with ASD as well as those with more varied language ability, including those at risk for language impairment.

Current Study

The present study utilized an eye-gaze paradigm (looking-while-listening) to assess real-time lexical processing in order to examine whether toddlers with ASD differ from neurotypical children in their use of ME during lexical processing. A major benefit of this paradigms is that it has fewer task demands compared to other popular paradigms and therefore allow greater inclusion of children with ASD from across the entire spectrum (Venker & Kover, 2015). In order to address the gap in the current literature in understanding the use of ME in toddlers on the autism spectrum, the present study employed an eye-tracking task which tapped this constraint in a lexical processing task involving referent selection (i.e., the tendency to look to an object when named). In this study the toddlers with ASD were matched to NT toddlers on nonverbal cognition. The current study was designed to answer two research questions. First, do toddlers with ASD differ from neurotypical children (matched on nonverbal cognition) in their use of ME during lexical processing (i.e., referent selection)? Second, is language ability related to use of ME?

Methods

Participants

The final matched sample consisted of 32 children with autism (24–36 months; 10 females) and 26 younger, neurotypical controls (18–24 months; 10 females), see Table 1. The children in the ASD group had significantly lower receptive language abilities than the children in the NT group. Due to this task's inquiry into a cognitive mechanism, we were interested in maintaining variability in the language abilities of the participants, while comparing children with ASD to NT children with similar cognitive abilities. Groups were matched ($p = .505$) on their visual reception raw score on the *Mullen Scales of Early Learning* (MSEL; Mullen, 1995). This difference in age was expected, allowing for a broader range of language and cognitive abilities to be included within the ASD group. We selected nonverbal cognition as the matching variable in order to include an ASD sample with a wider range of language and cognitive abilities than has been done in prior research. A subset of children with ASD have concomitant intellectual disability; it is possible to include these children by matching on nonverbal cognitive scores and comparing ASD toddlers with lower cognition to younger controls with average range cognition. Participants in the ASD group received a DSM-5 ASD diagnosis from an experienced psychologist during the visit (see Standardized Assessments). Exclusionary criteria for both groups included uncorrected hearing or vision impairments, known chromosomal abnormalities, cerebral palsy, fetal alcohol syndrome, seizure disorders, or other known neurological disorders. Additionally, children in the NT group were excluded if there were signs of developmental delay per parent report on a background information form, standardized assessment scores or behavior observed by the team psychologist and speech-language pathologist, or if the child was at an increased risk for ASD (i.e., scored above the 0-2 'low risk of ASD' range) based on a parent-report autism screening, the *Modified Checklist for Autism in Children* (M-CHAT; Robins et al. 2001).

[Table 1 here]

Procedure

All participants took part in a two-day visit which included developmental testing and several experimental eye-gaze tasks. Children with autism also participated in a full autism diagnostic protocol (for full protocol details see Standardized Assessments). Information on other eye-gaze experiments conducted within this larger project can be found in previous publications (e.g., Ellis Wesimer et al. 2016; Mahr et al. 2015; Venker et al., 2019; Pomper et al., 2019). Children were seen for an initial visit between 1.5 – 2 years of age for the NT group and 2-3 years of age for the ASD group. Each visit lasted approximately one hour for the NT children, and 2.5 hours for the children with ASD (to allow time for the in-depth parent interview and other diagnostic testing). Participants were recruited through local early intervention programs, pediatricians, and a research registry within the Waisman Center at the University of Wisconsin-Madison. Parents of children in both groups provided written consent for participation based on the protocol approved by the Education and Social/Behavioral Science Institutional Review Board at the University of Wisconsin-Madison.

Standardized Assessments

Several developmental, cognitive and language assessments were administered across the two-day research visit. Children in the ASD group participated in the *Autism Diagnostic Observation Schedule, 2nd Edition,* ADOS-2 (Lord et al., 2012) and their parents participated in the *Autism Diagnostic Interview- Revised*, ADI-R (Rutter, 2003), administered by an experienced (research reliable) psychologist, to diagnose ASD and provide a measure of autism severity. Children received Module 1, 2 or the Toddler Module of the ADOS-2, based on age and language ability. Children in both the ASD and NT groups participated in the Visual Reception and Fine Motor scales of the *Mullen Scales of Early Learning* (MSEL; Mullen, 1995) administered by a licensed psychologist, in order to assess nonverbal cognitive abilities. Parents completed the MacArthur Bates Communicative Development Inventories (MB-CDI) Words and Gestures and Words and Sentences subscales. The Auditory Comprehension and Expressive Communication scales of the *Preschool Language Scales, 5th Edition* (PLS-5; Zimmerman et al. 2011), were administered by a certified speech language pathologist to children in both groups, to assess receptive language and expressive language abilities.

Experimental Task

Half of the children participated in the Mutual Exclusivity task on the first day of their two-day visit and the other half were administered this task on the second day. Task administration was randomized to avoid order effects. Toddlers were seated on their caregiver's lap in front of a 55-inch wall-mounted television screen. Children were told they were going to watch a video but were given no explicit instructions. Caregivers' eyes were covered by opaque glasses in order to prevent them from unintentionally influencing the child's responses.

Caregivers were instructed not to talk to their children nor to direct their attention in any way during the experiment. A video camera mounted below the television screen recorded the child's face during the experiment for later offline eye-gaze coding. Caregivers were instructed that they could stop the task at any time for any reason. Overall, toddlers appeared to maintain interest and attention to the task.

Stimuli

Auditory stimuli consisted of a carrier phrase and target label (e.g., "Look at the shoe!") followed by a reinforcing phrase (e.g., "That's cool!). Nouns chosen were those that were likely to be known by toddlers with ASD based on local norms from a prior sample of toddlers with ASD (N=129) at approximately 30 months (Ellis Weismer, unpublished data). Nouns were also confirmed to be known by most of the children in both the ASD and NT groups via parent report from the MB-CDI. Presentation of auditory stimuli was consistent across trials to allow for direct comparison, such that the onset of the target noun occurred at the same point in each trial. Stimuli were 12 yoked object pairs of 12 real word nouns (*truck, blanket, ball, book, chair, bed, duck, hat, cup, apple, shoe, slide*) and 12 nonword nouns (*dax, dofa, jick, boge, fisp, tever, feeb, neidge, pum, shan, lort, pafe*). Objects were depicted in images on grey backgrounds presented against a black screen. Real word images were prototypical color photographs obtained via internet searches. Novel object images and labels were obtained through the NOUN data base (Horst & Hout, 2016). Object pairs were counterbalanced such that each object was presented as both the target and the distractor. Two different task orders were created to ensure that results were not driven by the order in which images were presented. There were two trial types within the task. In both trial types, a pair of images (one familiar object and one novel object) was presented. Images were presented in silence for 1500 ms in order to allow for children to view both objects, followed by the presentation of a carrier phrase plus label (e.g., “Find the bed”) and reinforcing phrase (e.g., “That’s cool!”). In real word trials, the familiar object was named. In the nonword condition, the novel object was named with a nonword label (see Figure 1 for sample stimuli). Trial types were semi-randomly presented throughout the experiment, with no one trial type occurring for more than two trials in a row. Each of 12 image pairs was presented twice, with the novel object labeled once and the real object labeled once, for a total of 24 trials. Trials were 5 seconds long. Short, engaging videos (5 sec.) accompanied by music were presented approximately every 4 trials in order to maintain children’s attention.

[Figure 1 here]

Eye Gaze Coding and Processing

A video recording of the child’s face during the experiment was coded offline by trained coders. Coders used a standard protocol as established in previous studies employing this paradigm (Fernald, et al., 2008). Every 33 ms, the coder marked the location of the child’s eyes as on the left image, right image, or neither image (i.e., in between images or off screen). Coders were unaware of which object was the target and which was the distractor. One out of every five videos (20%) was coded by a second coder for inter-rater reliability. Reliability across all time frames was 98%. Reliability across time frames with a shift in gaze was 97%. The experimental window for this task was 500ms-1800ms after noun onset. This time window was selected because it is similar to previous work in which an empirical window was selected in order to contain the average rise and plateau of looks to the target (Venker et al., 2019; Barr, 2008). Trials in which children looked at the images for less than half of the critical window were excluded. Additionally, participants were excluded if they contributed four or fewer trials. Two participants were excluded from the final analyses due to contributing four of fewer trials of useable data, one from the ASD group and one from the NT group. After data cleaning, children in the ASD group contributed an average of 9.94 trials in the nonword condition and 10.4 trials in the real word condition, and children in the NT group contributed an average of 10.8 trials in the nonword condition and 10.0 trials in the real word condition.

Analysis Plan

To answer the first research question, growth curve analysis (Mirman, 2014) was used to model changes in the probability of looking to the target versus nontarget image over the course of the experimental window. We focused our analyses on the experimental condition (Nonword trials) as we were interested in children's referent selection of the nonword (i.e., use of ME) rather than their real word recognition. The dependent variable was the empirical log-odds of looks to the target versus the nontarget object. Consistent with other LWL studies, (Fernald et al., 2008) looks coded as neither on the target nor distractor image (i.e., off screen) were not included in the analyses. The independent variable was time, which was quantified using the same orthogonal polynomial time terms as in our previous research (Pomper et al., 2019). These time terms can be interpreted as follows: Intercept represents the average empirical log-odds across the entire window. Linear time represents the average slope of the line, which indicates rate of change in fixation proportion (i.e., efficiency). Quadratic time represents the rate of the symmetric rise and fall around the peak asymptote (i.e., accuracy) of fixation proportions. Cubic time captures the slope of the tails of the curve, therefore quantifying any delay in increased fixations to the target in response to the auditory cue. Group was included as a between-subjects factor and contrast coded. We included the interactions of Group and all three time terms[1]. We included age and receptive language ability (PLS-5 Auditory Comprehension Raw Score) (mean-centered) as covariates. We also included children's accuracy on real word trials as a covariate, to statistically account for children's knowledge of the nouns. To answer our second research question, we conducted additional growth curve analyses, including interacting effects for Auditory Comprehension Raw Scores from the PLS-5, collapsing across groups, as well as a median split analysis to examine differences between children with ASD and higher language ability compared to those with lower language ability. We chose to use receptive language scores because of the nature of the experimental task as a language processing, rather than production task. All linear mixed effect models included the full random effects structure with by-participant random intercepts and random slopes for condition and time terms (linear, quadratic, cubic).

Results

When collapsing across groups, there was a significant effect of the intercept, group, age, linear and cubic time (see Table 2). This indicates that children spent more time fixating the target than the distractor (i.e., intercept effect). But most importantly, children's accuracy is significantly higher at the end of the window compared to the beginning (linear effect). The cubic effect captures the asymptote at the beginning and end of the window. Children in the ASD and NT groups performed significantly differently in the experimental condition (see Figure 2). There was a significant effect of group on linear time, $b = 1.2$, $t = 1.97$, $p < .05$. This indicates that after the onset of the target novel word, the rate of increase in children's fixations to the novel object (i.e., from the beginning to the end of the critical window) was greater for children in the NT compared to ASD group. There were no significant effects of group on any of the other times terms, p's $> .08$ (see Table 2). The linear time effect is particularly important in this analysis, since both groups were looking to target above chance at baseline, likely due to novelty of the nonword targets. The effect of linear time indicates whether looks to target *increased* after the target was named.

[Table 2 and Figure 2 here]

We were also interested in the relationship between extant receptive language abilities and use of ME in this task for children with ASD and NT. We were particularly interested in this relationship within the ASD group due to the variability in structural language ability within this population (Anderson, et al., 2007; Tager-Flusberg & Kasari, 2013; Arunachalam & Luyster, 2016). To explore this relationship, we conducted a second set of growth curve analyses, adding receptive language (as measured by raw scores on the PLS-5 Auditory Comprehension subscale). Receptive language was mean centered in the model. When receptive language ability was added to the model and allowed to interact with the polynomial time terms, there were no significant effects of group on any time terms, p's > .23, but there was a significant effect of receptive language on linear time $b = 0.20$, $t = 3.16$, $p < .01$ (see Table 3). This suggests that, collapsing across groups, when receptive language is accounted for, there are no longer group differences on any time terms. However, children with higher receptive language ability are demonstrating increased looks to target after the naming of the noun than those with lower language ability. Moreover, this may suggest that receptive language ability moderates the relationship between group and linear time which was observed in our first analysis.

[Table 3 here].

To explore this interaction further within the ASD group, we conducted a series of median split analyses based on PLS-5 Auditory Comprehension Raw Scores. The ASD group was split into a higher language group and lower language group about the median of the full sample (n = 18 lower language, n = 14 higher language). A within-group analysis revealed significant effects of intercept and linear time, significant interactions between language group and linear time, $b = 1.79$, $t = 2.45$, $p < .01$, and quadratic time, $b = 0.89$, $t = 2.16$, $p < .05$ (see Table 4), suggesting that children in the ASD group who had higher receptive language abilities demonstrated significantly different accuracy and efficiency of processing in the experimental (Nonword) condition than those with lower receptive language abilities (see Figure 3).

[Table 4 and Figure 3 here]

Discussion

The first aim of this study was to determine whether children with ASD differed from NT controls in their use of ME during lexical processing that entailed referent selection. This study provides the first evidence to date that, as a group, toddlers with ASD differed significantly from cognitively matched NT peers in their use of ME during a lexical processing task, when accounting for age and comprehension of familiar words. Additional analyses demonstrated that extant receptive language ability appears to play an important role in understanding these observed group differences. Previous studies have found that older children with ASD demonstrate comparable use of ME to NT peers matched on vocabulary (deMarchena et al., 2011; Hartley et al., 2019). Our findings suggest that when toddlers with ASD and lower receptive language abilities are included in the sample, resulting in a sample that is inclusive of children on the autism spectrum who may also show other associated deficits, such as language delay (Wiggins et al., 2015), those children with ASD and lower language ability indeed demonstrate less robust use of ME than NT peers. When evaluated as a whole group, toddlers with ASD demonstrated less robust ability to utilize this lexical constraint to infer that a novel label maps to a novel object during word processing, compared to NT controls, matched on cognitive ability. After the onset of the novel word, the increase in children's fixations to the novel object were significantly greater for children in the NT group compared to children in the ASD group. This difference is important to understand, as it may contribute to downstream delays, particularly for children with ASD who may also present with structural language disorder later in development. In addition to investigating a younger sample than in previous work, a strength of this study was that our sample of children with ASD included a broader range of language abilities than is often seen in studies that include a neurotypical control group. While our NT and ASD groups were matched on nonverbal cognition, the inclusion of varied language abilities allowed for examination of the use of ME across a broader portion of the autism spectrum, rather than only the subset of the children with ASD who do not present with language delays, which is often the case in prior work on ME (e.g., deMarchena et al., 2011). Importantly, when language ability was added to the model, there were no longer significant effects of diagnostic group (though this nonsignificant group effect was marginal), but there were significant interactions between receptive language ability and accuracy and efficiency of processing. This suggests that receptive language ability may moderate the relationship between diagnostic status and use of ME. Importantly, median split analyses revealed that when the ASD group was divided into higher and lower receptive language groups, significant differences in ME performance were revealed between the two language subsets such that those with ASD and higher language demonstrated more robust use of ME than those with ASD and lower language. These findings capture the significant heterogeneity in language outcomes for children with ASD and suggest that those children with lower receptive language abilities may struggle more with this constraint than children with relatively higher receptive language ability. Our findings suggest that there may be an important link between the ability to use the ME constraint and extant receptive language ability, particularly for children with ASD. While toddlers with ASD demonstrated less robust use of ME in this task than NT peers overall, this appears to be less driven by ASD diagnosis, and more by delayed language ability within a subset of the sample.

Fully understanding the differences in performance shown by the ASD group will require further inquiry. One area of interest is the potential impact of novelty bias. In general, toddlers prefer novel stimuli in ME tasks, particularly during referent selection (Horst et al., 2011; Kucker et al., 2018). Children in both groups appeared to have a bias toward the novel object at baseline, based on visualization of the data. This bias was not statistically evaluated. Still, an apparent novelty bias is notable, as a hyper-focus on the novel object may have prevented children from using their knowledge of the familiar object for the purposes of mutual exclusivity. That is, disambiguation of the novel object in this paradigm requires attention to both the novel the familiar objects in order to exclude the familiar object as a referent for the novel label (Hartley et al., 2019). Pomper and Saffran (2019) found that the salience of objects impacts NT toddlers' ability to use ME during referent selection. While the perceptual salience of the object pairs on each trial was designed to be balanced, it is possible that the unfamiliar nature of the novel objects made them inherently more salient.

The second aim of this study was to understand the relationship between extant language abilities and use of ME in both groups. Our findings are consistent with studies of older children with ASD (deMarchena et al., 2011; Hartley et al., 2019) which have suggested that some autistic children show evidence of access to mechanisms of language learning, such as ME, while others show delayed or impaired access to these mechanisms. Our findings suggest that those children who struggle with this mechanism also appear to have lower language ability, but we do not yet have evidence of the directionality of this relationship. These findings are important for understanding the possible implications of the ME constraint in this population, particularly for a subset of the ASD population at risk for concomitant language impairment. The association with extant language ability in the ASD group is consistent with the hypothesis that ME impacts language acquisition in these children. However, in this study we have only demonstrated an association between ME and language comprehension. Further research will be needed to determine whether there is indeed a causal link between ME deficits in young children with ASD and poor word learning outcomes. We do not know, as of yet, the directionality of the relationship between ME performance and language ability or word learning. It is possible that use of constraints such as ME helps children learn more words, or that knowing more words helps children engage in ME. Alternatively, the relationship may be bi-directional. Further research will be needed in order to disentangle this relationship. We were specifically interested in the relationship between use of ME and receptive language ability within the ASD group, as young children with ASD have demonstrated evidence of an uneven language profile, with receptive language abilities relatively more impaired than expressive language abilities (Davidson & Ellis Weismer, 2017; Hudry et al., 2010; Volden et al., 2011). However, future research might also investigate potential links between expressive language abilities and use of ME in this population.

While these results are meaningful in the context of referent selection, we acknowledge that this study does not answer the question of whether meaningful label-object pairings were mapped, learned, and retained over time. The retention and generalization of label-object pairings cannot be deduced from this study. Given that Horst & Samuelson (2008) found that fast-mapped label-object pairings are not always retained by toddlers over time, we cannot infer that potential word-object mappings using ME would have been retained by our participants in this task. Understanding the use of ME during referent selection is an important first step in gaining insight into the role of this cognitive mechanism in language development of toddlers with ASD. However, additional investigation into how ME facility is associated with word learning over time will be important to understanding the potential implications of these differences on vocabulary acquisition. While Hartley et al. (2019) have studied retention and generalization of ME in school-age children with ASD, future research is needed to investigate the generalization of novel object-label pairings using ME in toddlers with ASD. Additionally, while our study included a neurotypical control group, future studies could include a group of late talking toddlers (without ASD) to more directly examine the role of diagnostic status over and above language ability in driving differences in the use of ME.

This study provides the first evidence of differences in the use of ME between toddlers with ASD and neurotypical peers matched on nonverbal cognition. Given that our samples were matched on nonverbal cognition, and our findings (accounting for age) suggested that there were group differences when language ability was not accounted for, but these group differences were no longer significant when language ability was accounted for, these findings provide evidence for the importance of language in understanding the ME mechanism. Findings from our median split analyses, which revealed discrepant profiles of performance on the ME task between lower and higher language subsets of the ASD group, further emphasize the role of language ability in understanding these group differences. Moreover, these findings might lend further support to those theoretical accounts which view ME as a domain-specific linguistic mechanism, rather than one that is linked to cognitive abilities more generally. Consistent with prior studies testing school-aged children with ASD, our findings suggest a relationship between receptive language ability and the use of ME in toddlers with ASD. Median split analyses revealed that within the ASD group, there were significant differences in performance between children with higher and lower receptive language ability. These results advance our current understanding of early differences in the use of a language learning mechanism in toddlers with ASD, particularly for those children with lower language ability, who may have been excluded from prior studies. The implications of these differences will require further research in order to be fully understood. Further inquiry into the impact of various levels of salience (i.e., salience due to novelty, perceptual salience) on the use of ME in toddlers with ASD is warranted. Additionally, the impact of autism severity as well as cognitive factors such as attention and memory may help us to better understand the role of ME in language development.

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Washington, DC: American Psychiatric Association.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K., & Pickles, A. (2007). Patterns of Growth in Verbal Abilities Among Children With Autism Spectrum Disorder. *Journal of Consulting and Clinical Psychology*. https://doi.org/10.1037/0022-006X.75.4.594

Arunachalam, S., & Luyster, R. J. (2016). The Integrity of Lexical Acquisition Mechanisms in Autism Spectrum Disorders: A Research Review. *Autism Research*, *9*, 810–828. https://doi.org/10.1002/aur.1590

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

Bedford, R., Gliga, T., Frame, K., Hudry, K., Chandler, S., Johnson, M. H., & Charman, T. (2013). Failure to learn from feedback underlies word learning difficulties in toddlers at risk for autism. *Journal of Child Language*, *40*(1), 29–46. https://doi.org/10.1017/S0305000912000086

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language,* 17(2), 417–431.

Davidson, M.M., & Ellis Weismer, S. (2017). A Discrepancy in Comprehension and Production in Early Language Development in ASD: Is it Clinically Relevant? *Journal of Autism and Developmental Disorders, 47,* 2163-2175.

de Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, *119*(1), 96–113. https://doi.org/10.1016/j.cognition.2010.12.011

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37(5), 630–641.

Ellis Weismer, S., Haebig, E., Edwards, J., Saffran, J., & Venker, C. E. (2016). Lexical Processing in Toddlers with ASD: Does Weak Central Coherence Play a Role? *Journal of Autism and Developmental Disorders*, *46*(12), 3755–3769. https://doi.org/10.1007/s10803-016-2926-y

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. . Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics: On-line methods in children's language processing* (pp. 97–135).

Golinkoff, R., Mervis, C., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*(1), 125-155. doi:10.1017/S0305000900008692

Grassmann, S., Schulze, C., & Tomasello, M. (2015). Children's level of word knowledge predicts their exclusion of familiar objects as referents of novel words. *Frontiers in Psychology,* 6.

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*, *11*, 1251–1263. https://doi.org/10.1111/jcpp.12734

Halberda, J. (2003). The development of a word-learning strategy. *Cognition,* 87(1), B23–B34.

Hartley, C., Bird, L. A., & Monaghan, P. (2019). Investigating the relationship between fast mapping, retention, and generalisation of words in children with autism spectrum disorder and typical development. *Cognition*, *187*(February), 126–138. https://doi.org/10.1016/j.cognition.2019.03.001

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. https://doi.org/10.3758/s13428-015-0647-3

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157.

Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition*, *118*(2), 234–244. https://doi.org/10.1016/j.cognition.2010.10.015

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. International Journal of Language and Communication Disorders, 45(6), 681–690. doi:10.3109/13682820903461493.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too Much of a Good Thing: How Novelty Biases and Vocabulary Influence Known and Novel Referent Selection in 18-Month-Old Children and Associative Learning Models. *Cognitive Science*, *42*, 463–493. https://doi.org/10.1111/cogs.12610

Law, F. II & Edwards, J. R. (2015). Effects of Vocabulary Size on Online Lexical Processing by Preschoolers. *Language Learning and Development*, 11:4, 331-355.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2019). The role of developmental change and linguistic experience in the mutual exclusivity effect. *PsyArXiv Preprint*, (2013). https://doi.org/https://doi.org/10.31234/osf.io/wsx3a

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1–4. Torrence, CA: Western Psychological Services.

Mahr, T., McMillan, B. T. M., Saffran, J. R., Ellis Weismer, S., & Edwards, J. (2015). Anticipatory coarticulation facilitates word recognition in toddlers. *Cognition*, *142*, 345–350. https://doi.org/10.1016/j.cognition.2015.05.009

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77. https://doi.org/10.1016/0364-0213(90)90026-S

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*(3), 241–275. https://doi.org/10.1016/S0010-0285(03)00034-3

Mather, E., & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*, 14 (1), 60–76.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, 119(4), 831.

Mirman, D. (2014). Growth curve analysis and visualization using R. Boca Raton, FL: CRC Press.

Mullen, E. M. (1995). Mullen scales of early learning. Minneapolis, MN: AGS Edition.

Pomper, R., Ellis Weismer, S., Saffran, J., & Edwards, J. (2019). Specificity of Phonological Representations for Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(8), 3351–3363. https://doi.org/10.1007/s10803-019-04054-5

Pomper, R., & Saffran, J. R. (2019). Familiar Object Salience Affects Novel Word Learning. *Child Development*, *90*(2), e246–e262. https://doi.org/10.1111/cdev.13053

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, *97*(1), 13–23.

Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, 29(6), 819–865.

Rutter, M., LeCouteur, A., & Lord, C. (2003). Autism Diagnostic Interview-Revised. Los Angeles: Western Psychological Service.

Scott-Van Zeeland, A. A., McNealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biological Psychiatry*, *68*(4), 345–351. https://doi.org/10.1016/j.biopsych.2010.01.011

Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R., Lord, C., Messinger, D., Thurm, A., & Halladay, A. (2016). Prospective Longitudinal Studies of Infant Siblings of Children With Autism: Lessons Learned and Future Directions. *J Am Acad Child Adolesc Psychiatry*, *55*(3), 179–187. https://doi.org/10.1016/j.gde.2016.03.011

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

Thorup, E., Kleberg, J. L., & Falck-Ytter, T. (2017). Gaze Following in Children with Autism: Do High Interest Objects Boost Performance? *Journal of Autism and Developmental Disorders*, *47*(3), 626–635. https://doi.org/10.1007/s10803-016-2955-6

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, 10, 201–224.

Venker, C. E., Pomper, R., Mahr, T., Edwards, J., Saffran, J., & Ellis Weismer, S. (2019). Comparing automatic eye tracking and manual gaze coding methods in young children with autism spectrum disorder. *Autism Research*, 1–13. https://doi.org/10.1002/aur.2225

Venker, Courtney E. & Kover, S. T. (2015). An Open Conversation on Using Eye-Gaze Methods in Studies of Neurodevelopmental Disorders. *Journal of Speech, Language, and Hearing Research*, *24*(2), 1–14. https://doi.org/10.1044/2015

Volden, J., Smith, I., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., et al. (2011). Using the Preschool Language Scale, Fourth Edition to characterize language in preschoolers with autism spectrum disorders. American Journal of Speech-Language Pathology, 20(3), 200–208. doi:10.1044/1058-0360(2011/10-0035)

Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., Diguiseppi, C., … Marshalyn, T. (2015). Autism Spectrum Disorder Symptoms Among Children Enrolled in the Study to Explore Early Development (SEED). *Journal of Autism and Developmental Disorders*, *45*(10), 3183–3194. https://doi.org/10.1007/s10803-015-2476-8

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). Preschool language scales (5th ed.). San Antonio, TX: The Psychological Corporation.

Table 1. Participant characteristics for the group with autism spectrum disorder (ASD) and the neurotypical (NT) group, with effect sizes (Hedges' g) for group differences.

	ASD Group (*n* = 32) Mean (*SD*) Range	NT Group (*n* = 26) Mean (*SD*) Range	Group Difference *p* value Hedges' *g*
Chronological age (months)	30.69 (3.54)	20.38 (1.6)	p < .05
	24 - 36	18 - 23	3.63
Auditory Comprehension	64.03 (12.96)	103 (14.03)	p < .05
Standard Score	50 - 98	77 - 124	2.9
Expressive Communication	76.69 (10.71)	107.23 (9.99)	p < .05
Standard Score	50 - 100	91 - 130	2.94
Total Language	68.53 (10.76)	105.42 (11.66)	p < .05
Standard Score	50 - 95	86 - 129	3.3
Nonverbal Ratio IQ	74.5 (12.64)	109.58 (13.62)	p < .05
	52 -102	86 - 145	2.68
Visual Reception	25.34 (3.37)	25.96 (3.64)	p = .505
Raw Score	20 - 36	20 - 34	0.18
Fine Motor Raw Score	22.31 (2.31)	23.16 (3.07)	p = .235
	20 – 27	10 - 28	0.32
ASD Symptom Severity	8.28 (1.6)	--	--

Note. Auditory Comprehension and Expressive Communication were measured by the Preschool Language Scales, 5th Edition (PLS-5). Nonverbal Ratio IQ and Visual Reception was measured by the Mullen Scales of Early Learning (MSEL). ASD Symptom Severity was measured by the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2).

Table 2. Model containing nonword condition, age as a covariate, collapsing across NT and ASD groups.

	Estimate	SE	*t* value	*p* value
Intercept	0.605	0.084	7.204	< .001
Linear time	1.373	0.308	4.459	< .001
Quadratic time	0.094	0.181	0.521	.603
Cubic time	-0.283	0.127	-2.225	.026
Group	0.743	0.328	2.263	.024
Real Word Accuracy	-0.292	0.473	-0.618	.537
Receptive Language	-0.003	0.016	-0.166	.868
Age	0.053	0.025	2.094	.036
Linear time: Group	1.216	0.616	1.975	.048
Quadratic time: Group	0.124	0.363	0.341	.733
Cubic time: Group	-0.447	0.254	-1.759	.079

Note. The independent variable was time, and the dependent variable was the empirical log-odds of looks to the target versus the nontarget object. Real word accuracy was represented by by-subject average accuracy on Real Word trials, during the experimental window. Age in months and Receptive Language Ability, represented by PLS-5 Auditory Comprehension Raw Scores, were mean-centered in the model.

Table 3. Model containing nonword condition, age, and receptive language, collapsing across groups.

	Estimate	SE	*t* value	*p* value
Intercept	0.598	0.083	7.245	< .001
Linear time	1.318	0.284	4.634	< .001
Quadratic time	0.082	0.180	0.458	.647
Cubic time	-0.277	0.127	-2.192	.028
Group	0.608	0.331	1.838	.066
Receptive Language	0.023	0.018	1.268	.205
Real Word Accuracy	-0.293	0.473	-0.621	.535
Age	0.053	0.025	2.096	.036
Linear time: Group	0.195	0.653	0.299	.765
Quadratic time: Group	-0.099	0.413	-0.239	.811
Cubic time: Group	-0.345	0.291	-1.187	.235
Linear time: Receptive Language	0.196	0.062	3.161	.002
Quadratic time: Receptive Language	0.043	0.039	1.083	.279
Cubic time: Receptive Language	-0.020	0.028	-0.709	.478

Note. The independent variable was time, and the dependent variable was the empirical log-odds of looks to the target versus the nontarget object. Real word accuracy was represented by by-subject average accuracy on Real Word trials, during the experimental window. Receptive Language was represented by PLS-5 Auditory Comprehension Raw Scores and was mean-centered in the model. Age in months was mean-centered in the model.

Table 4. Model containing nonword condition, receptive language median split analysis within ASD group.

	Estimate	SE	*t* value	*p* value
Intercept	0.493	0.118	4.193	< .001
Linear time	0.853	0.366	2.329	.020
Quadratic time	0.074	0.206	0.361	.718
Cubic time	-0.072	0.158	-0.455	.649
Language Group	0.112	0.235	0.476	.634
Linear time: Language Group	1.794	0.732	2.451	.014
Quadratic time: Language Group	0.891	0.412	2.162	.031
Cubic time: Language Group	-0.262	0.316	-0.829	.407

Note. The independent variable was time, and the dependent variable was the empirical log-odds of looks to the target versus the nontarget object. Language groups represent a median split within the ASD group based on Preschool Language Scales, 5th Edition Auditory Comprehension Raw Scores. ASD group was split into a higher language group and lower language group about the median of the full sample (n = 18 lower language: n = 14 higher language).

Figure Captions

Figure 1 Sample visual and auditory stimuli. Top: Real word condition, auditory cue: "Find the shoe;" Bottom: Nonword condition, auditory cue "Find the jick"

Figure 2 *Raw looking behavior, full trial:* Time course is plotted for the full trial in the nonword condition (blue line) and real word condition (red line). Time courses are plotted separately for subsamples of the ASD and NT groups matched on nonverbal cognition (left and right columns, respectively). Lines are raw data best fit lines with ribbons representing ± 1 standard error of the mean.

Figure 3 *Growth curve analyses, ASD higher and lower language groups:* Time course is plotted for the experimental window (500 – 1800ms after noun onset) in the nonword condition. Data from the ASD lower language group (n = 18) and ASD higher language group (n = 14), are plotted. Fixations are plotted as the empirical log-odds. The line at log-odds 0 represents chance (i.e., equal looks to target vs. distractor object). Data points represent observed data averaged across participants. Lines are growth curve fits with ribbons representing ± 1 standard error of the mean. Median split was conducted based on Preschool Language Scales, 5th Edition Auditory Comprehension Raw Scores within the ASD group.

Footnotes

[1]We did not include the three-way interactions between Group, Receptive Language, and each time term. This omission was based on both conceptual reasons (the correlation between children's receptive language ability and their accuracy in spoken word recognition should not differ between groups) and methodological reasons (including this interaction would force the

model to extrapolate beyond the data, because a substantial number of children in the ASD group had PLS scores below the minimum for the NT group).