Exemplar Variability and Cross-Situational Word Learning in Young Autistic Children

By

Kathryn E. Prescott

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Communication Sciences and Disorders)

at the

UNIVERSITY OF WISCONSIN-MADISON
2024

Date of final oral examination: May 9, 2024

The dissertation is approved by the following members of the Final Oral Committee:

Susan Ellis Weismer, Professor Emeritus, Communication Sciences and Disorders

Haley Vlach, Associate Professor, Educational Psychology

Audra Sterling, Associate Professor, Communication Sciences and Disorders

Jenny Saffran, Professor, Psychology

Table of Contents

List of Tables

List of Figures

Dedication

For anyone trying to find the right words

Acknowledgements

I would first like to acknowledge the funding sources supporting this work, including NIH grants NIDCD F31 DC020901 (Prescott, PI), NIDCD R01 DC017974 (Ellis Weismer & Saffran, MPIs), and NICHD P50 HD105353 (Waisman Center core grant), as well as the Council of Academic Programs in Communication Sciences and Disorders (CAPCSD).

I will forever be grateful to my advisor, Susan Ellis Weismer, for taking a chance on me and believing in my potential to grow as a researcher. Your mentorship, support, and insight will continue to inspire me throughout my career. I sincerely thank you for the incredible contributions your knowledge and work have provided, not only to your students but to the entire field of child language disorders. It has been a tremendous privilege to work with you.

I would also like to thank my F31 co-sponsor, Haley Vlach, for shaping my research interests, sharing wisdom about life in and out of academia, and helping me think about language acquisition in new ways that will carry my work moving forward. I am also very grateful to Jenny Saffran whose passion, knowledge, and incomparable expertise in language sciences I give tremendous credit for inspiring my intellectual curiosity from day one of my doctoral program. I also thank Audra Sterling for the unending support she has generously provided through rigorous independent study and personal mentorship, teaching opportunities, welcoming me to her lab meetings, and so much more. To my entire committee, I am so grateful and humbled by all you have done in supporting my development as a scientist.

Thank you to the members of the Language Processes Lab for their support of this project, including our coders, our lab manager Kristine, and our clinical team Heidi, Martha, and Lucia. Special thanks to Janine and Carrie for their companionship, commiseration, and advice. Thanks to my fellow doctoral students in language sciences across CSD, Psychology, and

Educational Psychology for all the support, friendship, and intellectual discussions that have enriched my doctoral experience.

My profound gratitude to the families who participated in this study, my former speech-language therapy clients and colleagues, and the autistic community for informing the purpose and direction of my career.

Finally, I must acknowledge the endless love and support provided by my incredible family and friends both near and far. Especially Mom, Dad, Tom, and John – I could not have done any of this without you. To my wonderful husband Jayse, words simply cannot express how grateful I am to have had you by my side through this experience. You make every day a little brighter and every challenge a little easier. A special thanks to my brilliant mother-in-law Amy for generously offering her expert advice on statistical analyses. To my late grandmother, Maryjane, thank you for inspiring my love of language and life-long pursuit of learning. I will forever be working to live up to your example. Love you all so much.

Abstract

What cognitive mechanisms allow young autistic children to learn words amidst the complexity of natural language environments? This dissertation sought to inform this question by investigating the ability of 2- to 4-year-old autistic children and neurotypical (NT) peers to learn words by tracking statistical co-occurrences between labels and referents across multiple contexts (i.e., cross-situational learning (CSL). While previous CSL studies examined children's ability to map each novel word to a single object, in natural environments children must track the co-occurrences between novel words and the many objects to which that word may refer. Thus, to elucidate the extent to which autistic children can harness cross-situational statistics to learn words referring to perceptually variable objects, I also administered a CSL task that included three exemplars varying in color for each novel word. I further examined whether explicit naming would improve word learning from variable exemplars. Finally, I investigated child characteristics predictive of individual differences in word learning across these conditions among autistic children. Analyses revealed that while autistic children did not show evidence of learning in the basic CSL task, they did demonstrate learning in the CSL task with exemplar variability. Further, they outperformed younger NT children on the task when nonverbal cognition was covaried. Explicit naming of variable exemplars did not improve performance for the autistic children and improved performance in the NT children to a greater extent compared to autistic children. Individual differences analyses revealed receptive language was associated with autistic children's performance in the basic CSL task and chronological age was associated with performance when variable exemplars were explicitly named. In sum, results indicated that preschool-age autistic children as a group could successfully use cross-situational statistics to learn words with (but not without) exemplar variability, and that learning was not improved by explicit naming. However, within-group variability in word learning was evident across tasks, explained in part by individual differences in receptive language and chronological age. More research will be needed to fully understand the mechanisms underlying this surprising pattern of findings, which might include the extent to which visual attention is directed by the perceptual salience of stimuli. While preliminary, these findings may inform future research, clinical practice, and environmental accommodations tailored to the individual strengths and needs of young autistic children.

Chapter 1: Introduction

Delayed "first words" is often one of the earliest and most visible signs of developmental difference among autistic children (Mitchell et al., 2006; Nitzan et al., 2022).[1] However, structural language abilities can be highly heterogeneous across the autism spectrum as development unfolds (Tager-Flusberg & Joseph, 2003; Anderson et al., 2007; Wittke et al., 2017). While many autistic individuals understand and produce words within the average range for their age (e.g., Wittke et al., 2017), around 20-30% remain non-speaking into middle childhood (Anderson et al., 2007; Lord et al., 2004). Many autistic children experience early language delays, producing first words at an average age of 23 months compared to 12 months as 75% of neurotypical (NT) children do (Mayo et al., 2013; Schneider et al., 2015). Though early language development has been found to be delayed on average, there is considerable within-group variability (Ellis Weismer et al., 2010; Rescorla & Safyer, 2013). Several studies have reported null differences in lexical-semantic abilities between autistic and NT children. For example, autistic children can map labels to objects at basic and superordinate category levels (Tager-Flusberg, 1985) and generally do not make more mapping errors than expressive vocabulary-matched NT peers in developmentally appropriate word learning tasks, even when required to integrate social information (Luyster & Lord, 2009). By the age of school entry, many autistic children perform as well as same-age NT peers on vocabulary tests (Kelley et al., 2006; Kjelgaard & Tager-Flusberg, 2001; Wittke et al., 2017). Other studies have indicated that autistic children do not differ from mental age-matched NT peers when sorting objects into categories according to function, color, and physical form (Ungerer & Sigman, 1987) or to age- and cognition-matched NT children in identifying typical category exemplars (Gastgeb et al.,

---

[1] Throughout this dissertation, I intend to adhere to the terminology preferred by the autistic community as outlined by Monk et al. (2022).

2006). Autistic children also appear to share the tendency of NT peers to map novel words to objects as opposed to actions (e.g., the noun bias; Swensen et al., 2007; Tek et al., 2008). However, results differ across studies depending on group characteristics and the basis of comparison. While studies have not found differences in early lexical composition among autistic children, NT children, and Late Talkers matched on expressive language, (Ellis Weismer et al., 2011; Rescorla & Safyer, 2013), when preverbal and minimally-verbal autistic children were compared to NT peers matched on size of productive vocabulary, differences in lexical composition were detected (Haebig et al., 2021). Together, evidence suggests lexical-semantic abilities vary considerably in autistic children depending on the task and individual characteristics. However, even in those with quantitatively similar language abilities as same-age NT peers, subtle differences have been found at more abstract levels of the lexical-semantic domain – namely, categorization and lexical organization (see Naigles & Tek, 2017 for a review). In this chapter, I review the lexical-semantic phenotype in autism including lexical organization and acquisition mechanisms, describe the broad theoretical framework informing this work, and summarize the studies comprising this dissertation.

Lexical Organization and Categorization

The extant research on lexical-semantic organization and categorization can be broadly divided into two areas of study: those investigating the structure and organization of existing lexical-semantic knowledge, and those examining processing of new information. Within the former, evidence for diagnostic group differences is conflicting. A study of school-age children participating in a lexical decision task demonstrated that autistic children did not differ in accuracy and efficiency in responding to words from large and small semantic networks and nonwords compared to same-age NT children with similar receptive language and nonverbal cognition (Haebig et al., 2015). Studies examining existing within-category knowledge have found that autistic children produced either smaller (Pastor-Cerezuela et al., 2016) or larger (Begeer et al., 2014) clusters of related words on semantic fluency tasks in comparison to age- and cognition-matched NT children. Other studies found autistic children were comparable to NT peers in processing of typical exemplars of familiar categories but may process atypical exemplars more slowly (Gastgeb et al., 2006), or perform less stably overall (Ellawadi et al., 2017). Autistic children have also been found to produce less prototypical exemplars than NT children with similar age and nonverbal cognition (Foldager et al., 2022) or receptive language (Dunn et al., 1996) for some known categories. Semantic priming studies have also provided insight into existing semantic organization in autistic children, with similarly conflicting findings. Autistic school-age children demonstrated semantic priming effects in several studies, showing evidence of knowledge of the semantic relation between prime and target (Hala et al., 2007; Harper-Hill et al., 2014; Toichi & Kamio, 2001). However, other studies reported findings of mixed or smaller semantic priming effects in autistic children in comparison to NT peers (Henderson et al., 2011; Kamio et al., 2007). These conflicting findings are likely due to methodological differences. Henderson et al. (2011) reported no differences semantic priming between groups at the stage when processing was more automatic, only detecting weaker priming effects in the autistic group at later stages of processing. Meanwhile, Kamio et al. (2007) presented prime and target words orthographically but did not measure variability in reading skills, which could account for the diagnostic group differences reported.

Differences have also emerged in studies examining the formation of categorical representations or processing of new semantic information, though the evidence is likewise mixed. One view of category learning suggests NT children form categories (including lexical categories) on the basis of an abstracted “ideal” or composite of category exemplars, called a prototype (Molesworth et al., 2008). New exemplars are then judged against this prototype to determine category membership. Differences in prototype formation has been proposed as one possible mechanism underlying differences in lexical-semantic abilities in autism. For example, in one study, school-age autistic children as a group did not identify a novel prototypical exemplar (the mathematical average of all previously presented category exemplars) as a lexical category member, unlike receptive vocabulary-matched NT children (Klinger & Dawson, 2001). Findings of differences in prototype formation and use compared to age-matched NT children with similar cognitive abilities have also been reported in the perception literature (Church et al., 2010; Gastgeb et al., 2009). However, other studies of autistic children with language and cognitive abilities in the average range for their age contradicted this finding, suggested that prototype learning may be more heterogeneous across the autism spectrum in relation to other abilities (i.e., generalization ability, nonverbal cognition) or depend on elements of task design (i.e., task ambiguity linguistic demands; Church et al., 2015; Molesworth et al., 2005, 2008). Autistic children have also demonstrated differences in categorical induction (Gelman & Markman, 1986). Studies have reported that despite having vocabulary scores in the average range, autistic children had more difficulty than NT peers extending category properties to novel objects given the same label (i.e., if a white rabbit eats grass, so should a brown rabbit) (Kelley et al., 2006) with residual difficulties persisting into adolescence (Naigles et al., 2013). It is important to note that categorization and generalization difficulties are likely also heterogeneous across the autism spectrum depending on individual characteristics and developmental level and are not limited to vocabulary learning. McGregor and Bean (2012) found that autistic children understood that nouns label nested, hierarchical categories of objects, but those with weaker structural language ability were less able to use social context to constrain category boundaries than age-matched NT peers, resulting in overly broad categories. A study of categorical induction in autistic adolescents and adults with average range cognition and language did not find differences from same-age NT comparisons (Soulières et al., 2011) suggesting developmental change in this ability among the subset of autistic individuals represented by this sample. In the perceptual literature, autistic adults demonstrated no differences in perceptual categorization, but poorer generalization compared to age and nonverbal cognition-matched NT adults (Froehlich et al., 2012).

Word Learning Biases

Several studies have also investigated NT attentional word learning biases as candidate processes underlying developmental differences in language in autism. English-speaking NT children employ strategies such as the whole object principle (assuming labels refer to an entire object rather than its constituent parts; Markman, 1990), the noun bias (mapping novel words to objects rather than actions; Gentner, 1982), mutual exclusivity (mapping novel words to novel objects rather than objects with known labels; Markman & Wachtel, 1988), and the shape bias (extending novel words to novel objects based on object shape rather than other perceptual features such as size, color, or texture; Landau et al., 1988). These processes aid word learning by constraining the otherwise infinitely possible word-to-object mappings present in the learning environment (Quine, 1960). Autistic children, however, show mixed evidence of reliance on the same word learning biases as NT children. Studies suggest autistic children do have a noun bias, tending to map novel words to objects rather than actions (Swensen et al., 2007; Tek et al., 2008). Several studies have indicated that school-age autistic children demonstrate mutual exclusivity during referent (de Marchena et al., 2011; Hartley et al., 2019; Preissler & Carey,

2005) while a study on mutual exclusivity in autistic preschoolers suggests use of this strategy may be attenuated compared to nonverbal cognition-matched NT peers (Mathée-Scott et al., 2022). Among these findings, one word learning constraint stands out as the most consistent evidence of differences in autism - the shape bias (Potrzeba et al., 2015; Tek et al., 2008).

Shape Bias

The shape bias refers to the tendency to generalize novel words to novel objects based on object shape rather than other features such as texture, color, or size, as first observed by Landau, Smith, and Jones (1988). In the classic shape bias paradigm, children are presented with a novel object labeled with a novel word "this is a *dax*." They are then presented with an array of objects matching the *dax* on only one dimension and asked "Which of these is also a *dax*?" For example, one object matches the *dax* on shape (but not color or size), while the other two objects would match the *dax* on another dimension (e.g., color or size) but not shape. Both NT adults and children as young as two years old consistently choose the novel object matching the *dax* on shape, extending the novel word and classifying objects based on shape rather than size or texture (Landau et al., 1988). Smith and colleagues (Smith, 1999; Smith et al., 2002; Smith & Samuelson, 2006) have proposed a mechanistic theory explaining this phenomenon in NT children, known as the Attentional Learning Account (ALA). According to the ALA, early word learning is acquired through simple associative learning processes. As these first words are learned, the structure of the input begins to teach an association between the syntax used for count nouns and object shape. Thus, the structure of children's early input trains their attention to relevant object category features (i.e., shape for early count nouns), in turn training future learning on the basis of shape. Indeed, corpus analyses have suggested that children's early lexicons contain the regularities necessary to build a shape bias (Samuelson & Smith, 1999). In further support of this account, Smith and colleagues have demonstrated that the shape bias emerges after children have acquired ~50 count nouns in their productive vocabulary, typically around 2 years of age for NT children (Landau et al., 1988; Smith et al., 2002). In contrast to the ALA, others have argued that the shape bias reflects underlying conceptual knowledge about object features and kinds, which may be present far earlier than language production (Booth et al., 2005; Booth & Waxman, 2008; Gelman, 2003). These researchers have argued that conceptual information may constrain word learning from birth in addition to perceptual and linguistic information.

As a group, young autistic children have not generally demonstrated a shape bias in previous work (Hartley & Allen, 2014; Potrzeba et al., 2015; Tek et al., 2008; Tovar et al., 2020). In both a longitudinal group comparison study and extension study by the same research group, NT children showed a consistent shape bias from 20-24 months onward, while language-matched autistic children as a group did not have a shape bias at any visit (Potrzeba et al., 2015; Tek et al., 2008). In a study with school-age children, "low-functioning"[2] autistic children were found to generalize novel labels to objects based on color *and* shape, shape only or color only, while language-matched NT children generalized based only on shape (Hartley & Allen, 2014), while Tovar et al. (2020) found that autistic children tended to extend object labels based on color and novelty. Another study of school-age children concluded that autistic children with low "verbal mental age" do not show a shape bias, while children with high verbal mental age (>4;6) do when an object is named (Field et al., 2016). The authors interpret this finding as suggesting that shape bias development may be delayed (but not absent) in autism. Indeed, some individual children – particularly those with larger vocabularies – do show a shape bias, and vocabulary is

---

[2] This term is potentially offensive and non-preferred by the autism community (Monk et al., 2022) but is included in this description due to its use in the cited paper.

related to shape bias performance in autism (Potrzeba et al., 2015). However, Potrzeba and colleages (2015) did not find any threshold effect of count or shape nouns for shape bias development among autistic children. Moreover, one study reported a negative correlation between known shape nouns and shape bias in autistic children with high language ability (but not low language ability), though the relationship did not reach significance in regression models (Abdelaziz et al., 2018). These findings suggest that vocabulary acquisition may not be the key mechanism underlying shape bias development in autism, as has been proposed for NT children (Samuelson & Smith, 1999; Smith et al., 2002) and Late Talkers (Jones, 2003). Preschoolers with specific language impairment (SLI; now known as a subtype of developmental language disorder; DLD) also do not tend to demonstrate a shape bias (Collisson et al., 2015). Given the role of vocabulary in shape bias development among other children with and without language impairments (LI; Collisson et al., 2015; Jones, 2003; Perry et al., 2022) and the heterogenous findings in autism (Potrzeba et al., 2015; Abdelaziz et al., 2018), it is likely that learning regularities from vocabulary does support the development of a shape bias (and in turn future lexical development) for at least some autistic children. However, since vocabulary delay does not appear to fully explain this difference, researchers have proposed alternative mechanisms. One possibility is that early conceptual difficulties in autism underlie differences in shape bias development (Booth et al., 2005; Booth & Waxman, 2008; Tek & Naigles, 2017). As described by Tek and Naigles (2017), autistic children may have difficulties with the conceptual foundations that underlie word learning, which could include abstraction processes necessary for category formation and induction. Possible cognitive mechanisms have also been explored using computational modeling (Tovar et al., 2020). In this study, an atypical shape bias in autism was modeled by local hyperplasticity and hyperconnectivity. Hyperplasticity and hyperconnectivity of the local processing networks resulted in fragmented categorical representations (rather than a generalized prototype representing the abstraction of previous learning) which compete to attract processing of incoming stimuli, further preventing robust, well-organized categories. The computational modeling was further supported by behavioral data demonstrating a tendency among autistic children to extend novel labels to objects based on novelty, regardless of perceptual properties (such as shape; Tovar et al., 2020). Outside of the laboratory, an early lack of shape bias may have cascading consequences for early count noun learning. For example, a recent semantic analysis of large language corpora revealed that the expressive lexical profile of minimally- and pre-verbal autistic toddlers (producing 1-10 words) differed from younger, expressive vocab-matched NT, containing proportionally more verbs and fewer nouns (Haebig et al., 2021). While their semantic knowledge and organization may differ from NT peers, autistic children clearly learn words. Empirical findings suggest one process through which autistic children may acquire new words is by tracking statistical associations between labels and referents across different contexts, known as cross-situational learning.

Cross-Situational Learning

Statistical learning theory posits that learners are sensitive to statistical regularities in the input. This ability enables infants as young as 8 months old to extract statistical regularities in verbal input after only 2 minutes exposure (Saffran et al., 1996). In terms of word learning, specifically label-referent mapping, children's sensitivity to the statistical correlations between labels and their referents facilitates word learning even with high levels of referential uncertainty, when no other associative cues are present. Previous evidence suggests that children and adults are able to learn the meanings of words simply through their repeated co-occurrence with a referent across contexts – referred to as cross-situational learning (CSL; Smith et al., 2011). In the classic CSL paradigm, learners are exposed to repeated trials in which two or more novel referents are presented with two or more spoken novel words, with no additional cues provided. Across trials, each word-referent pair occurs repeatedly with a variety of other word-referent pairs. As such, the only information available to learners to disambiguate the word-referent pairings is their statistical co-occurrence across these varying contexts over multiple exposures. Using cross-situational statistics, learners can acquire adjectives (Akhtar & Montague, 1999), verbs (Scott & Fisher, 2012), nouns (Vlach & Johnson, 2013; Yu & Smith, 2007), and noun categories (Chen et al., 2018). CSL can even allow for simultaneous learning of word-referent pairings from multiple grammatical categories (Rebuschat et al., 2021). Though CSL performance may improve with age over the preschool years (Aktar & Montague, 1999), children as young as 12 months can quickly acquire novel word-object pairs presented together across contexts (Smith & Yu, 2008). How do young children accomplish this feat? One account claims that simple associative learning mechanisms could be responsible (Yu, 2008; Yu et al., 2007). Two possible strategies using associative learning could underlie CSL: A frequentist approach, wherein associative links are strengthened or weakened as co-occurrences or non-co-occurrences accumulate, or an elimination approach, involving formulation and evaluation of hypotheses. However, researchers have posited no fundamental mechanistic difference between these two strategies as both rely on associative learning (Kachergis et al., 2012). It is also possible that CSL reflects a continuum of learning strategies underpinned by simple associative mechanisms depending on the level of referential uncertainty in the task. When referential uncertainty is high, a frequentist approach may be most beneficial, while an elimination approach may be more appropriate for low referential uncertainty (Smith et al., 2011). Regardless of the specific strategy employed, evidence supports the idea that CSL is essentially associative in nature and relies on aggregating information across instances because there are effects of contextual diversity (Suanda et al., 2014). That is, children's learning was improved with more variability in the learning set, indicating that learning depends on co-occurrence information accrued across contexts.

CSL has also been investigated in school-age children with developmental language disorder (DLD). Across studies, children with DLD were outperformed by peers with typical language abilities (Ahufinger et al., 2021; Broedelet et al., 2023; McGregor et al., 2022) though they showed word learning accuracy above chance (Ahufinger et al., 2021; Broedelet et al., 2023) and learned words at a similar rate to peers with typical development (McGregor et al., 2022). Interestingly, one study reported that children with DLD did not rely on statistical aggregation as a learning strategy in a CSL task and their learning outcomes were not related to extant vocabulary. These findings suggest children with DLD may therefore employ different mechanisms to learn words than peers with typical language (McGregor et al., 2022).

Autistic children have demonstrated relative strengths in behavioral studies of statistical learning across a variety of tasks (e.g., serial reaction time, contextual cuing, word segmentation, and artificial grammar learning; see Obeid et al., 2016 for a meta-analysis; but see Hu et al., 2023 for contradictory findings). Regarding cross-situational word learning specifically, three studies to date have investigated this ability in school-age and adolescent autistic children. Children in this age group (5-9 years, Venker, 2019; 5-15 years, Hartley et al., 2020), learned words with accuracy that did not differ from younger, receptive vocabulary-matched NT peers on a CSL task. Moreover, children did not differ in the CSL task compared to ostensive teaching, where objects are explicitly labeled in a 1:1 naming context (Hartley et al., 2020; Venker, 2019). However, Hartley et al. (2020) did report that autistic children were significantly slower to identify named referents. In the third study, autistic children (*M* age = 134 months) also did not differ from NT peers of a similar age and nonverbal cognitive ability on a CSL task, again performing better than chance on novel word-object mapping (McGregor et al., 2013). However, autistic children across studies varied considerably within group. Some children learned well, while others showed little evidence of learning (Venker, 2019). Learning was associated with language measures, including familiar word processing (Venker, 2019) and receptive vocabulary (McGregor et al., 2013), but not level of autism traits or nonverbal cognition (Venker, 2019). Accuracy of referent selection during training (but not retention or generalization of learning) was also associated with receptive vocabulary (Hartley et al., 2020). To date, no studies have yet examined CSL in autistic children before the age of school entry, so it is not known whether autistic children in the early stages of language acquisition are able to track statistical structure in the input to learn object-label mappings. Another unknown, even for older autistic individuals, is whether word learning is differentially affected by noise in the input surrounding the statistical structure – in other words, variability.

Learning from Variability

While not yet examined in the linguistic domain for autistic children, research going back decades emphasizes the role of variability in learning (Estes & Burke, 1953). Variability decontextualizes representations, thereby improving generalization (Schmidt & Bjork, 1992; Twomey et al., 2018). Variability in exemplars may encourage comparison among category members and draw attention to common structure, thus informing categorization (e.g., Namy & Gentner, 1999). Without variability, early lexical representations may be over-specified (Apfelbaum & McMurray, 2011; Sandhofer & Schonberg, 2020). However, not all variability is helpful. A training set is more effective when variability occurs on dimensions that are irrelevant to category structure – otherwise, children might associate noncontrastive cues with objects (Apfelbaum & McMurray, 2011). Theoretically, simple associative mechanisms may also account for the role of variability in word learning (Apfelbaum & McMurray, 2011). From the lens of the Attentional Learning Account (Smith & Samuelson., 2006), within-category variability informs attentional biases by drawing attention to the underlying structure of the category - and over time, the structure of the input broadly - to form specific attentional biases such as the shape bias (Sandhofer & Schonberg, 2020). In terms of the statistical learning framework, variability of irrelevant features increases the salience of more statistically consistent, relevant structures, facilitating their learning (Aguilar et al., 2018; Gomez, 2002; Plante et al., 2014).

Extant evidence broadly supports the facilitative role of variability in children's learning of novel words for concrete objects. In one study, 18-month-old children were taught novel nominal categories with either variable (dissimilar) or highly similar exemplars. All children learned trained labels, but variable exemplars facilitated generalization to novel exemplars and the development of a more specified word learning bias. That is, children who learned from variable exemplars learned not only when to apply the shape bias, but also when *not* to apply a shape bias, resulting in faster vocabulary gains (Perry et al., 2010). Another study taught 2-year-olds a novel noun category through multiple referent selection trials where a novel object was presented in the context of familiar objects and given a novel label. Children were placed in one of two conditions. In the single exemplar condition, children saw the same exemplar throughout training. In the multiple exemplar condition, children were presented with several within-category exemplars throughout training, differing in color but matching in shape. Children in both conditions did well with immediate referent-selection, but only those in the exemplar variability condition retained the novel noun category after a delay (Twomey et al., 2014). In a second experiment, the researchers compared children's novel noun learning between narrow within-category variability – where the multiple exemplars differed only in color - and wide within-category variability, where exemplars differed across multiple dimensions, such as color, texture, and size. Only children in the narrow variability condition retained learning after a delay (Twomey et al., 2014). Variation across too many perceptual dimensions in the wide within-category variability condition likely increased the difficulty of initial category formation. However, the authors did not account for minor modifications in shape among within-category exemplars stimuli in their wide variability stimuli set that were not present in the narrow variability stimuli, which may have been particularly detrimental to word learning for children in the wide variability condition. This finding emphasizes the importance of shared structure (usually shape for concrete nouns) across within-category exemplars for variability to facilitate word learning. Aside from variation in perceptual features across exemplars, other forms of variability may also impact upon children's word learning. For example, variability in the background on which novel objects are presented across trials facilitated retention of label-object mappings in 2-year-old children (Twomey et al., 2018). The authors suggested that background variability acts as a decontextualization mechanism for lexical representations, better specifying the essential structure of the representation by increasing the signal-to-noise ratio.

However, input variability may not be helpful for all learners in all contexts. For example, in a study of 4- to 7-year-old bilingual and monolingual NT children, speaker variability overall did not affect immediate novel label-object mapping in a CSL task compared to listening to only one speaker (Crespo & Kaushanskaya, 2021). However, children with lower sustained attention performed better in the speaker variability condition. This finding may lend support to attention as a mechanistic pathway through which variability improves word learning. In another study with a similar sample (5- to 8-year-old monolingual and bilingual children), speaker variability was combined with exemplar variability to determine whether children learned as well with variability across multiple modalities, as is often required in naturalistic learning. In this study, generalization to a new exemplar was the proxy measure for word learning. Results indicated that neither exemplar variability nor speaker variability alone conferred word learning advantages over no variability, and that combined exemplar-speaker variability did not differ from variability in a single modality for children overall. However, bilingual children were more accurate than monolingual children in the combined variability condition (Crespo et al., 2023). It also appears that very young learners may not benefit from multiple exemplars during word learning. Taxitari et al. (2020) found that 10-month-old NT infants were able to learn word-object mappings from a single exemplar but failed to learn when presented with multiple exemplars during training. The researchers attributed this result to excessive cognitive demands for such young learners. Another study presenting contrary evidence found that 2.5- and 3-year-old NT children demonstrated better novel verb learning and extension when viewing only one actor repeatedly compared to four different actors (Maguire et al., 2008). Given that mapping a novel label to an action may be more challenging than mapping to a novel object (e.g., Gentner, 1982, 2006; Gentner & Boroditsky, 2001), these findings may suggest that variability is less beneficial if the difficulty of a task is already high. In sum, the extant literature suggests that while variability appears to be broadly beneficial to word learning in NT children (at least for concrete nouns), the effect may not be evident in all children to the same extent. Might children with neurodevelopmental disorders, particularly those with language difficulties, also benefit from variability in the input?

Though data are limited, it appears the answer is yes for children with developmental language disorder (DLD; sometimes referred to in older literature as specific language impairment; language impairment, or language learning disabilities; Bishop et al., 2017). In one study, 4- to 5-year-old children with DLD were taught novel words corresponding to either three different exemplars or a single exemplar. While children in both conditions mapped and generalized the novel words when tested immediately after training, those in the exemplar variability condition retained that learning significantly better after a 3-week delay (Aguilar et al., 2018). Another study examined the effect of high or low variability in verbs used in a conversational recast treatment for preschool children with language impairment. Results suggested that high variability of unique verbs led to gains in trained (but not untrained) morphemes (Plante et al., 2014). In addition, more children in the high-variability condition showed a strong treatment effect and produced more unique utterances containing trained morphemes compared to the low-variability condition. A final study including individuals with impaired language investigated the effect of visual variability in the training of academic concepts to college students. In the high-variability condition, participants were as accurate on novel items as trained items, while participants in the low-variability condition were less accurate on novel items. Students with language learning difficulties showed the same pattern of performance as students with typical language abilities but were less accurate overall (Bourgoyne & Alt, 2017).

To date, no studies have investigated the effect of variability in the word learning of autistic individuals. However, evidence from nonlinguistic domains suggests variability may not be as useful to at least a subset of autistic children. In a perceptual category learning task, 7-12-year-old autistic children who were considered “high-functioning” but had difficulty with generalization learned better from prototype-only training than from perceptually-variable exemplars, while those with generalization skills similar to NT individuals generalized better when trained with variable exemplars (Church et al., 2015). The researchers noted that a large portion of the autistic participants had difficulty with generalization, even with a sample consisting only of "high-functioning" individuals with language and cognitive abilities in the average range for their age. If children with a wider range of abilities had been included, it is likely that atypical generalization (and thus poorer learning from variability) would have been evident in an even larger portion of the autistic group.

Summary and Theoretical Framework

In sum, the lexical-semantic profile of autism that has emerged from previous literature can be characterized by heterogeneity across individuals, studies, and specific processes, but also by some broad trends. On average, autistic children have demonstrated early language delays despite considerable within-group variability (Ellis Weismer et al., 2010; Rescorla & Safyer, 2013). Depending on group characteristics and matching criteria, at least a subset of autistic children have been found not to differ from NT peers on many language measures, including standardized vocabulary assessments (Kelley et al., 2006; Kjelgaard & Tager-Flusberg, 2001; McGregor et al., 2012), label-object mapping (Luyster & Lord, 2009; Tager-Flusberg, 1985), mapping novel words to objects rather than actions (Swensen et al., 2007; Tek et al., 2008), and several linguistic statistical learning tasks (Haebig et al., 2017; Mayo & Eigsti, 2012; see Obeid et al., 2016 for a meta-analysis, but see Hu et al., 2023). Cross-situational learning (CSL) studies have yielded no differences in fast mapping and generalization of novel words between autistic children and receptive vocabulary-matched NT peers (Venker, 2019; Hartley et al., 2020). However, the youngest autistic children included in previous studies were 4- to 5-years old, leaving the developmental emergence of CSL an open question to be addressed in this dissertation. Autistic children have diverged from NT lexical-semantic development in previous literature more often in lexical organization and categorization, particularly on tasks involving greater abstraction and complexity (i.e., Church et al., 2015; Gastgeb et al., 2006; Kelley et al., 2006; Naigles et al., 2013). While autistic children as a group do appear to have a noun bias (Swensen et al., 2007; Tek et al., 2008), evidence is more mixed for mutual exclusivity (de Marchena et al., 2011; Hartley et al., 2019; Mathée-Scott et al., 2022; Priessler & Carrey, 2005). As a group, autistic children generally do not tend to demonstrate a shape bias though development may be delayed (Field et al., 2016; Hartley & Allen, 2014; Potrzeba et al., 2015; Tek et al., 2008). However, several questions remain: can autistic children use cross-situational statistics to acquire words in ambiguous contexts during the early stages of language development? Does variability affect word learning for autistic children, particularly in light of possible differences in underlying mechanisms such as abstraction as may be evidenced by shape bias studies? What characterizes individual differences in word learning across the autism spectrum?

The studies comprising this dissertation sought to answer these questions within a framework informed by a convergence of statistical learning theory (e.g., Saffran, 1996) and the Attentional Learning Account (e.g., Smith, 1999; Smith & Samuelson, 2006). That is, implicit simple associative learning mechanisms may allow children to begin detecting statistical regularities among elements in their environment, including co-occurrences between words and their referents across contexts over time, even in the absence of any additional cues (e.g., CSL). Then, as children learn the statistical structure of their linguistic environment, they may use that structural knowledge toward future learning by attending more to relevant perceptual and linguistic features of current input (Smith & Samuelson, 2006). Variability in the input aids this process by rendering the invariant structure of lexical categories more salient (Aguilar et al., 2018; Gomez, 2002; Plante et al., 2014). As children continue to encounter variable exemplars given the same label, their lexical representations become increasingly robust and abstract, supporting generalization of learning (see Sandhofer & Schonberg, 2020 for a review).

Past work with autistic school-age children and adolescents suggests that CSL does not differ on average from NT children with similar receptive vocabulary (Venker, 2019; Hartley et al., 2020). However, it is not yet known if this ability comes online as early in development for autistic children as NT peers. Even in NT children, developmental changes in CSL abilities have been observed across the preschool years (Akhtar & Montague, 1999; Vlach & DeBrock, 2019). Moreover, early differences in visual attention (Elsabbagh et al., 2013; Keehn et al., 2013; Mo et al., 2019; Sacrey et al., 2013) could constrain even simple associative learning of labels and their referents and result in early language delays, warranting investigation of CSL in younger autistic children. Therefore, the first aim of this dissertation will be to examine the early development of CSL in autistic children and compare their learning to NT peers.

Even if associative learning is not broadly affected by diagnostic status, the broad lack of shape bias in early autistic development noted in previous literature suggests these children may experience differences in lexical-semantic category formation or organization, which may be at least in part attributed to underlying processes such as abstraction of relevant categorical properties (Naigles & Tek, 2017). Tovar and colleagues (2020) have offered a more detailed mechanistic explanation derived from computational modeling: differences in shape bias result from an increased excitation/inhibition ratio resulting in hyperplasticity and hyperconnectivity, which in turn lead to over-specified and fragmented categorical representations (rather than an abstracted prototype) competing for processing of incoming stimuli and preventing the formation of robust, de-contextualized, and generalizable categories. Critically, previous CSL studies with autistic children have paired only a single object with each word, whereas learning a word in the real world requires children to form lexical categories, associating a single label with the shared category-relevant structure of many objects that vary across category-irrelevant perceptual features. If, as Tovar et al. (2020) propose, children's categorical representations are fragmented and highly context-bound, each new exemplar is perceived as highly novel (i.e., hyperplasticity), and abstraction is thus impaired, then one might expect autistic children to have difficulty learning in the context of exemplar variability and particularly in generalizing that learning to new instances. Thus, the second aim of this dissertation will be to investigate the effect of exemplar variability in young autistic children's word learning, both in referentially ambiguous (i.e., CSL) and ostensive teaching contexts.

Finally, in light of the heterogeneity inherent to the autism spectrum, the third aim of this dissertation was to explore individual child characteristics associated with word learning in referentially ambiguous (i.e., CSL), perceptually variable, and referentially ambiguous *and* perceptually variable contexts. The goal of this investigation will be to better explain the heterogeneity in word learning among autistic children and align with the principles of precision medicine in future language interventions, better specifying which word learning contexts are most appropriate for individual autistic children (Beversdorf, 2016; Loth et al., 2016).

Overview of Present Studies

Cross-Situational Word Learning in Young Autistic Children

The first study of this dissertation investigated the emergence of cross-situational word learning (CSL) in young autistic children in the early stages of language development. While older autistic children appear to be able to use cross-situational statistics to learn words in ambiguous contexts (Venker, 2019; Hartley et al., 2020), the prevalence of early language delays in autism could imply that CSL comes online at a later point in development relative to NT peers. To examine this possibility, 2- to 4-year-old autistic children and NT peers were administered an experimental eye-gaze task employing a similar paradigm to a previous CSL study (Venker, 2019). In each trial of the training phase of this task, children viewed images of two novel objects side-by-side on screen and heard two novel words, with no cues as to which word mapped to which object. The objects appeared in various combinations across trials but were always paired with the same novel word. Thus, the only way to successfully disambiguate the correct word-referent mappings is to track the statistical co-occurrences between objects and words across trials. This training was followed by a test phase where children would see two objects on screen in each trial (the target object and a distractor) and were asked "Where's the *dax*? Do you see it?" While no study has yet investigated CSL in autistic children in this developmental stage, NT children have demonstrated developmental effects in CSL abilities through the preschool years (Ahktar & Montague, 1999; Vlach & DeBrock, 2019) and young autistic children on average experience early delays in language acquisition (Howlin, 2003; Mayo et al., 2013). It is therefore possible that early developmental differences could constrain the ability to attend to cross-situational statistics in order to map words to referents. Therefore, while older autistic children have shown comparable CSL abilities to receptive language-matched peers in previous studies (Venker, 2019; Hartley et al., 2020; McGregor et al., 2013), 2- to 4-year-old children might experience relatively more difficulty using cross-situational statistics to learn novel words.

The Effect of Variable Exemplars in Young Autistic Children's Word Learning

The second study explored the impact of exemplar variability on novel noun learning in 3- to 4-year-old autistic children and NT peers. While no previous studies have directly investigated variability in the word learning of autistic children, evidence that autistic children on average do not demonstrate a shape bias could be suggestive of difficulties in abstraction of category-relevant features across multiple exemplars (Naigles & Tek, 2017). To explore this question, 3- to 4-year-old autistic and NT children participated in two experimental eye-gaze tasks involving CSL and ostensive teaching. For each novel word, children saw three different exemplars paired with it, all sharing the same shape but differing along a category-irrelevant perceptual property (color). The test trials followed the same structure as the CSL task in the first study. children saw two pictures side-by-side on screen (the target image and a distractor image) while hearing the target word spoken in the frame "Where's the *dax*? Do you see it?" Given diagnostic group differences on shape bias tasks in previous studies, I expected the autistic group to demonstrate more difficulty learning novel words from variable exemplars compared to a single exemplar, and more difficulty in the exemplar variability condition compared to NT peers. I also expected the autistic group to experience more difficulty with ambiguous labeling in the context of exemplar variability, demonstrating relatively better performance in the ostensive teaching task than the CSL task.

Individual Child Characteristics Associated with Autistic Children's Word Learning

The aim of the final study in this dissertation was to investigate individual child characteristics associated with autistic children's word learning in the experimental eye-gaze tasks of the first two studies. The purpose of this study was two-fold: First, to better understand individual differences underlying the heterogeneity in word learning across the autism spectrum under various conditions; second, to align this work with the principles of precision medicine (Beversdorf, 2016; Loth et al., 2016) with the aim of tailoring future language interventions to individual needs. To do so, children were administered a battery of autism diagnostic, language, and cognitive assessments compatible with remote assessment (due to the COVID-19 pandemic). I chose measures of receptive language, level of autism traits, and nonverbal cognition from these assessments as well as age in months to quantify individual child characteristics. The measure for receptive language was raw scores on the Auditory Comprehension (AC) subtest of the Preschool Language Scales - Fifth Edition (PLS-5; Zimmerman et al., 2011). The measure for level of autism traits was the Childhood Autism Rating Scale – 2nd Edition (CARS-2; Schopler et al., 2010) total raw score. Finally, the measure for nonverbal cognition was the Developmental Assessment of Young Children – 2nd Edition (DAYC-2; Voress & Maddox, 2013) cognitive domain raw score. Age was measured in months. In previous studies, CSL task performance has been associated with aspects of receptive language in autistic children (receptive vocabulary and familiar word processing, Venker, 2019; receptive vocabulary, Hartley et al., 2020; McGregor et al., 2013). Therefore, I expected PLS-5 AC raw scores to be positively related to word learning in the CSL tasks. The effect of exemplar variability has not yet been examined in autistic children, but if shape bias is associated with diagnostic status, one might expect level of autism traits to predict learning from exemplar variability.

References

Abdelaziz, A., Kover, S., Wagner, M., & Naigles, L. (2018). The Shape Bias in Children With Autism Spectrum Disorder: Potential Sources of Individual Differences. *Journal of Speech, Language, and Hearing Research*, *61*, 2685–2702. https://doi.org/10.1177/1461444810365020

Aguilar, J. M., Plante, E., & Sandoval, M. (2018). Exemplar variability facilitates retention of word learning by children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *49*(1), 72–84. https://doi.org/10.1044/2017_LSHSS-17-0031

Ahufinger, N., Guerra, E., Ferinu, L., Andreu, L., & Sanz-Torrent, M. (2021). Cross-situational statistical learning in children with developmental language disorder. *Language, Cognition and Neuroscience*, *36*(9), 1180–1200. https://doi.org/10.1080/23273798.2021.1922723

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347–358. https://doi.org/10.1177/014272379901905703

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K., & Pickles, A. (2007). Patterns of Growth in Verbal Abilities Among Children With Autism Spectrum Disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. https://doi.org/10.1037/0022-006X.75.4.594

Apfelbaum, K. S., & McMurray, B. (2011). Using variability to guide dimensional weighting: Associative mechanisms in early word learning. *Cognitive Science*, *35*(6), 1105–1138. https://doi.org/10.1111/j.1551-6709.2011.01181.x

Begeer, S., Wierda, M., Scheeren, A. M., Teunisse, J.-P., Koot, H. M., & Geurts, H. M. (2014). Verbal fluency in children with autism spectrum disorders: Clustering and switching strategies. *Autism*, *18*(8), 1014–1018. https://doi.org/10.1177/1362361313500381

Beversdorf, D. Q. (2016). Phenotyping, Etiological Factors, and Biomarkers: Toward Precision Medicine in Autism Spectrum Disorders. *Journal of Developmental and Behavioral Pediatrics*, *37*(8), 659–673. https://doi.org/10.1097/DBP.0000000000000351

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Consortium, and the C.-2. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Booth, A. E., & Waxman, S. R. (2008). Shape bias special section: Taking stock as theories of word learning take shape. *Developmental Science*, *11*(2), 185–194. https://doi.org/10.1111/j.1467-7687.2007.00664.x

Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental Psychology*, *41*(3), 491–505. https://doi.org/10.1037/0012-1649.41.3.491

Bourgoyne, A., & Alt, M. (2017). The effect of visual variability on the learning of academic concepts. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1568–1576. https://doi.org/10.1044/2017_JSLHR-L-16-0271

Broedelet, I., Boersma, P., & Rispens, J. (2023). Implicit cross-situational word learning in children with and without developmental language disorder and its relation to lexical-semantic knowledge. *Frontiers in Communication*, *8*. https://doi.org/10.3389/fcomm.2023.1021654

Chen, C., Zhang, Y., & Yu, C. (2018). Learning Object Names at Different Hierarchical Levels Using Cross-Situational Statistics. *Cognitive Science*, *42*(S2), 591–605. https://doi.org/10.1111/cogs.12516

Church, B. A., Krauss, M. S., Lopata, C., Toomey, J. A., Thomeer, M. L., Coutinho, M. V., Volker, M. A., & Mercado, E. (2010). Atypical categorization in children with high-functioning autism spectrum disorder. *Psychonomic Bulletin & Review*, *17*(6), 862–868. https://doi.org/10.3758/PBR.17.6.862

Church, B. A., Rice, C. L., Dovgopoly, A., Lopata, C. J., Thomeer, M. L., Nelson, A., & Mercado, E. (2015). Learning, plasticity, and atypical generalization in children with autism. *Psychonomic Bulletin and Review*, *22*(5), 1342–1348. https://doi.org/10.3758/s13423-014-0797-9

Collisson, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2015). Individual differences in the shape bias in preschool children with specific language impairment and typical language development: Theoretical and clinical implications. *Developmental Science*, *18*(3), 373–388. https://doi.org/10.1111/desc.12219

Crespo, K., & Kaushanskaya, M. (2021). Is 10 Better than 1? The Effect of Speaker Variability on Children's Cross-Situational Word Learning. *Language Learning and Development*, *17*(4), 397–410. https://doi.org/10.1080/15475441.2021.1906680

Crespo, K., Vlach, H., & Kaushanskaya, M. (2023). The effects of bilingualism on children's cross-situational word learning under different variability conditions. *Journal of Experimental Child Psychology*, *229*, 105621. https://doi.org/10.1016/j.jecp.2022.105621 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

Dunn, M., Gomes, H., & Sebastian, M. J. (1996). Prototypicality of responses of autistic, language disordered, and normal children in a word fluency task. *Child Neuropsychology*, *2*(2), 99–108. https://doi.org/10.1080/09297049608401355

Ellawadi, A. B., Fein, D., & Naigles, L. R. (2017). Category structure and processing in 6-year-old children with autism: Categorization in 6-year-old with Autism. *Autism Research*, *10*(2), 327–336. https://doi.org/10.1002/aur.1652

Ellis Weismer, S., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and grammatical skills in toddlers on the autism spectrum compared to late talking toddlers. *Journal of Autism and Developmental Disorders*, *41*(8), 1065–1075. https://doi.org/10.1007/s10803-010-1134-4

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

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of Visual Attention in Infancy is Associated with Emerging Autism in Toddlerhood. *Biological Psychiatry*, *74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030

Estes, W. K., & Burke, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, *60*(4), 276–286. https://doi.org/10.1037/h0055775

Field, C., Allen, M. L., & Lewis, C. (2016). Attentional learning helps language acquisition take shape for atypically developing children, not just children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *46*, 3195-3206.

Foldager, M., Vestergaard, M., Lassen, J., Petersen, L. S., Oranje, B., Aggernaes, B., & Simonsen, E. (2022). Atypical Semantic Fluency and Recall in Children and Adolescents with Autism Spectrum Disorders Associated with Autism Symptoms and Adaptive Functioning. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05677-x

Froehlich, A. L., Anderson, J. S., Bigler, E. D., Miller, J. S., Lange, N. T., DuBray, M. B., Cooperrider, J. R., Cariello, A., Nielsen, J. A., & Lainhart, J. E. (2012). Intact prototype formation but impaired generalization in autism. *Research in Autism Spectrum Disorders*, *6*(2), 921–930. https://doi.org/10.1016/j.rasd.2011.12.006

Gastgeb, H. Z., Rump, K. M., Best, C. A., Minshew, N. J., & Strauss, M. S. (2009). Prototype formation in autism: Can individuals with autism abstract facial prototypes? *Autism Research*, *2*(5), 279–284. https://doi.org/10.1002/aur.93

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do individuals with autism process categories differently? The effect of typicality and development. *Child Development*, *77*(6), 1717–1729. https://doi.org/10.1111/j.1467-8624.2006.00969.x

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford University Press.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209. https://doi.org/10.1016/0010-0277(86)90034-X

Gentner, D. (1982). *Why Nouns are Learned Before Verbs: Linguistic Relativity Versus Natural Partitioning. Technical Report No. 257.*

Gentner, D. (2006). Why Verbs Are Hard to Learn. In K. A. Hirsh-Pasek & R. M. Golinkoff, *Action Meets Word* (1st ed., pp. 544–564). Oxford University PressNew York. https://doi.org/10.1093/acprof:oso/9780195170009.003.0022

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (1st ed., pp. 215–256). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.010

Gomez, R. L. (2002). Variability and Detection of Invariant Structure. *Psychological Science*, *13*(5), 431–436.

Haebig, E., Jiménez, E., Cox, C. R., & Hills, T. T. (2021). Characterizing the early vocabulary profiles of preverbal and minimally verbal children with autism spectrum disorder. *Autism*, *25*(4), 958–970. https://doi.org/10.1177/1362361320973799

Haebig, E., Kaushanskaya, M., & Ellis Weismer, S. (2015). Lexical Processing in School-Age Children with Autism Spectrum Disorder and Children with Specific Language Impairment: The Role of Semantics. *Journal of Autism and Developmental Disorders*, *45*(12), 4109–4123. https://doi.org/10.1007/s10803-015-2534-2

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

Hala, S., Pexman, P. M., & Glenwright, M. (2007). Priming the Meaning of Homographs in Typically Developing Children and Children with Autism. *Journal of Autism and Developmental Disorders*, *37*(2), 329–340. https://doi.org/10.1007/s10803-006-0162-6

Harper-Hill, K., Copland, D., & Arnott, W. (2014). Efficiency of Lexical Access in Children with Autism Spectrum Disorders: Does Modality Matter? *Journal of Autism and Developmental Disorders*, *44*(8), 1819–1832. https://doi.org/10.1007/s10803-014-2055-4

Hartley, C., & Allen, M. L. (2014). Brief report: Generalisation of word-picture relations in children with autism and typically developing children. *Journal of Autism and Developmental Disorders*, *44*(8), 2064–2071. https://doi.org/10.1007/s10803-014-2074-1

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*(March), 126–138. https://doi.org/10.1016/j.cognition.2019.03.001

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*(July 2019), 104265. https://doi.org/10.1016/j.cognition.2020.104265

Henderson, L. m., Clarke, P. j., & Snowling, M. j. (2011). Accessing and selecting word meaning in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *52*(9), 964–973. https://doi.org/10.1111/j.1469-7610.2011.02393.x

Howlin, P. (2003). Outcome in High-Functioning Adults with Autism with and Without Early Language Delays: Implications for the Differentiation Between Autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, *33*(1).

Hu, A., Kozloff, V., Owen Van Horne, A., Chugani, D., & Qi, Z. (2023). Dissociation Between Linguistic and Nonlinguistic Statistical Learning in Children with Autism. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-05902-1

Jones, S. S. (2003). Late talkers show no shape bias in a novel name extension task. *Developmental Science*, *6*(5), 477–483. https://doi.org/10.1111/1467-7687.00304

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychonomic Bulletin and Review*, *19*(2), 317–324. https://doi.org/10.3758/s13423-011-0194-6

Kamio, Y., Robins, D., Kelley, E., Swainson, B., & Fein, D. (2007). Atypical Lexical/Semantic Processing in High-Functioning Autism Spectrum Disorders without Early Language Delay. *Journal of Autism and Developmental Disorders*, *37*(6), 1116–1122. https://doi.org/10.1007/s10803-006-0254-3

Keehn, B., Müller, R.-A., & Townsend, J. (2013). Atypical attentional networks and the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *37*(2), 164–183. https://doi.org/10.1016/j.neubiorev.2012.11.014

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders*, *36*(6), 807–828. https://doi.org/10.1007/s10803-006-0111-4

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). For Genetic Subgroups. *Language and Cognitive Processes*, *16*(2–3), 287–308.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*(1), 111–124. https://doi.org/10.1017/S0954579401001080

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Lord, C., Risi, S., & Pickles, A. (2004). Trajectory of language development in autistic spectrum disorders. In *Developmental language disorders* (pp. 18–41). Psychology Press.

Loth, E., Murphy, D. G., & Spooren, W. (2016). Defining Precision Medicine Approaches to Autism Spectrum Disorders: Concepts and Challenges. *Frontiers in Psychiatry*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyt.2016.00188

Luyster, R. J., & Lord, C. (2009). Word Learning in Children with Autism Spectrum Disorders. *Developmental Psychology*, *45*(6), 1774–1786. https://doi.org/10.1037/a0016223.Word

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*(4), 628–634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

Markman, E. M. (1990). Constraints Children Place on Word Meanings. *Cognitive Science*, *14*(1), 57–77. https://doi.org/10.1207/s15516709cog1401_4

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

Mathée-Scott, J., Larson, C., Venker, C., Pomper, R., Edwards, J., Saffran, J., & Ellis Weismer, S. (2022). Use of Mutual Exclusivity and its Relationship to Language Ability in Toddlers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *52*(10), 4528–4539. https://doi.org/10.1007/s10803-021-05321-0

Mayo, J., Chlebowski, C., Fein, D. A., & Eigsti, I.-M. (2013). Age of First Words Predicts Cognitive Ability and Adaptive Skills in Children with ASD. *Journal of Autism and Developmental Disorders*, *43*(2), 253–264. https://doi.org/10.1007/s10803-012-1558-0

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

McGregor, K. K., & Bean, A. (2012). How Children With Autism Extend New Words. *Journal of Speech, Language, and Hearing Research*, *55*(1), 70–83. https://doi.org/10.1044/1092-4388(2011/11-0024)

McGregor, K. K., Rost, G., Arenas, R., Farris-Trimble, A., & Stiles, D. (2013). Children with ASD can use gaze in support of word recognition and learning. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(7), 745–753. https://doi.org/10.1111/jcpp.12073

McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbisi-Kelm, T., & Pomper, R. (2022). What Children with Developmental Language Disorder Teach Us About Cross-Situational Word Learning. *Cognitive Science*, *46*(2). https://doi.org/10.1111/cogs.13094

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics*, *27*(2 SUPPL. 2). https://doi.org/10.1097/00004703-200604002-00004

Mo, S., Liang, L., Bardikoff, N., & Sabbagh, M. A. (2019). Shifting visual attention to social and non-social stimuli in Autism Spectrum Disorders. *Research in Autism Spectrum Disorders*, *65*, 56–64. https://doi.org/10.1016/j.rasd.2019.05.006

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2005). The prototype effect in recognition memory: Intact in autism? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*(6), 661–672. https://doi.org/10.1111/j.1469-7610.2004.00383.x

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2008). When prototypes are not best: Judgments made by children with autism. *Journal of Autism and Developmental Disorders*, *38*(9), 1721–1730. https://doi.org/10.1007/s10803-008-0557-7

Monk, R., Whitehouse, A. J. O., & Waddington, H. (2022). The use of language in autism research. *Trends in Neurosciences*, 1–3. https://doi.org/10.1016/j.tins.2022.08.009

Naigles, L. R., Kelley, E., Troyb, E., & Fein, D. (2013). Residual difficulties with categorical induction in children with a history of autism. *Journal of Autism and Developmental Disorders*, *43*(9), 2048–2061. https://doi.org/10.1007/s10803-012-1754-y

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), 1–12. https://doi.org/10.1002/wcs.1438

Namy, L. L., & Gentner, D. (1999). Comparison in the Development of Categories. *Cognitive Development*, *14*(4), 487–513.

Nitzan, T., Koller, J., Ilan, M., Faroy, M., Michaelovski, A., Menashe, I., Meiri, G., & Dinstein, I. (2022). The Importance of Language Delays as an Early Indicator of Subsequent ASD

Diagnosis in Public Healthcare Settings. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05757-y

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology*, *7*(AUG), 1–18. https://doi.org/10.3389/fpsyg.2016.01245

Pastor-Cerezuela, G., Fernández-Andrés, M.-I., Feo-Álvarez, M., & González-Sala, F. (2016). Semantic Verbal Fluency in Children with and without Autism Spectrum Disorder: Relationship with Chronological Age and IQ. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00921

Perry, L. K., Horst, S. C., Kucker, J. S., & Samuelson, L. K. (2022). Late Bloomer or Language Disorder? Differences in Toddler Vocabulary Composition Associated with Long-term Language Outcomes. *Developmental Science*. https://doi.org/10.1111/desc.13342

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189

Plante, E., Oglivie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A. M., & Burton, R. (2014). Variability in the Language Input to Children Enhances Learning in a Treatment Context. *American Journal of Speech-Language Pathology*, *23*, 530–545. https://doi.org/10.1044/2014

Potrzeba, E., Fein, D., & Naigles, L. (2015). Investigating the shape bias in typically developing children and children with autism spectrum disorders. *Frontiers in Psychology*, *6*(MAR), 1–12. https://doi.org/10.3389/fpsyg.2015.00446

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

Quine, W. V. O. (1960). *Word and Object*. Technology Press of the Massachusetts Institute of Technology.

Rebuschat, P., Monaghan, P., & Schoetensack, C. (2021). Learning vocabulary and grammar from cross-situational statistics. *Cognition*, *206*(November 2020), 104475. https://doi.org/10.1016/j.cognition.2020.104475

Rescorla, L., & Safyer, P. (2013). Lexical Composition in Children with Autism Spectrum Disorder (ASD). *Journal of Child Language*, *40*(1), 47–68. https://doi.org/10.1017/S0305000912000232

Sacrey, L. A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural Brain Research*, *256*, 441–450. https://doi.org/10.1016/j.bbr.2013.08.028

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition*, *73*(1), 1–33.

Sandhofer, C. M., & Schonberg, C. (2020). Multiple Exemplars Support Children's Word Learning: The Roles of Aggregation, Decontextualization, and Memory Dynamics. In *Language and Concept Acquisition from Infancy Through Childhood* (pp. 159–178). https://doi.org/10.1007/978-3-030-35594-4

Schmidt, R. A., & Bjork, R. A. (1992). New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychological Science*, *3*(4), 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x

Schneider, R. M., Yurovsky, D., & Frank, M. C. (2015). Large-scale investigations of variability in children's first words. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society, Pasadena, California, July 22–25, 2015*, 2110–2115.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale—Second Edition (CARS-2)*. Western Psychological Services.

Scott, R. M., & Fisher, C. (2012). 2.5-Year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*(2), 163–180. https://doi.org/10.1016/j.cognition.2011.10.010

Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science*, *35*(3), 480–498. https://doi.org/10.1111/j.1551-6709.2010.01158.x

Smith, L. B. (1999). Children's noun learning: How general learning processes make specialized learning mechanisms. In *The emergence of language* (pp. 277–303).

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19. https://doi.org/10.1111/1467-9280.00403

Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to cimpian and markman (2005) and booth, Waxman, and Huang (2005). *Developmental Psychology*, *42*(6), 1339–1343. https://doi.org/10.1037/0012-1649.42.6.1339

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Soulières, I., Mottron, L., Giguère, G., & Larochelle, S. (2011). Category induction in autism: Slower, perhaps different, but certainly possible. *Quarterly Journal of Experimental Psychology*, *64*(2), 311–327. https://doi.org/10.1080/17470218.2010.492994

Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology*, *126*, 395–411. https://doi.org/10.1016/j.jecp.2014.06.003

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development*, *78*(2), 542–557. https://doi.org/10.1111/j.1467-8624.2007.01022.x

Tager-Flusberg, H. (1985). Basic level and superordinate level categorization by autistic, mentally retarded, and normal children. *Journal of Experimental Child Psychology*, *40*(3), 450–469. https://doi.org/10.1016/0022-0965(85)90077-3

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

Taxitari, L., Twomey, K. E., Westermann, G., & Mani, N. (2020). The Limits of Infants' Early Word Learning. *Language Learning and Development*, *16*(1), 1–21. https://doi.org/10.1080/15475441.2019.1670184

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, *1*(4), 208–222. https://doi.org/10.1002/aur.38

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

Toichi, M., & Kamio, Y. (2001). Verbal Association for Simple Common Words in High-Functioning Autism. *Journal of Autism and Developmental Disorders*, *31*(5), 483–490.

Tovar, Á. E., Rodríguez-Granados, A., & Arias-Trejo, N. (2020). Atypical shape bias and categorization in autism: Evidence from children and computational simulations. *Developmental Science*, *23*(2), 1–14. https://doi.org/10.1111/desc.12885

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the Right Noises: Background Variability Helps Early Word Learning. *Cognitive Science*, *42*, 413–438. https://doi.org/10.1111/cogs.12539

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's More Like It: Multiple Exemplars Facilitate Word Learning. *Infant and Child Development*, *23*, 105–122. https://doi.org/10.1002/icd

Ungerer, J. A., & Sigman, M. (1987). Categorization skills and receptive language development in autistic children. *Journal of Autism and Developmental Disorders*, *17*(1), 3–16. https://doi.org/10.1007/BF01487256

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*(November 2018), 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *45*(4), 700–711. https://doi.org/10.1037/xlm0000611.Statistics

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015.Memory

Voress, J. K., & Maddox, T. (2013). *Developmental Assessment of Young Children-Second Edition (DAYC-2)*. PRO-ED.

Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., & Naigles, L. R. (2017). Grammatical Language Impairment in Autism Spectrum Disorder: Exploring Language Phenotypes Beyond Standardized Testing. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00532

Yu, C. (2008). A Statistical Associative Account of Vocabulary Growth in Early Word Learning. *Language Learning and Development*, *4*(1), 32–62. https://doi.org/10.1080/15475440701739353

Yu, C., & Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., Smith, L. B., Klein, K. A., & Shiffrin, R. M. (2007). Hypothesis Testing and Associative Learning in Cross-Situational Word Learning: Are They One and the Same? *Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)*.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Chapter 2: General Methods

Participants

Participants in the following studies were recruited as part of a large, longitudinal project examining prediction and language in young autistic children. Across all studies reported in this dissertation, 20 autistic children between 24 and 47 months old and 30 neurotypical (NT) children between 18 and 47 months old participated in and had data coded and analyzed for the final version of the experimental tasks. Four additional NT participants were tested but not coded or analyzed due to time constraints. Participant characteristics for children included in each study can be found in their respective chapters. The University of Wisconsin-Madison Institutional Review Board approved the research protocol. Participants were recruited from the local Madison community and surrounding region. Legal guardians for all participants provided written informed consent prior to study enrollment. Across both groups, children had no known hearing or vision impairments and were from monolingual English-speaking households. Children in the NT group demonstrated typical development across domains, including cognition and language. They displayed no features of autism, developmental delays, motor deficits, or seizures. Children were included in the autistic group if they met criteria for DSM-5 autism diagnosis based on administration of standardized autism diagnostic assessments (see Standardized Measures) and best clinical estimate by an experienced clinical psychologist on our research team. In the larger project and the present studies, children were included in the autistic group with concomitant cognitive and/or language delay to better represent the range of ability inherent to the autism spectrum (Anderson et al., 2007; Tager-Flusberg & Joseph, 2003; Wittke et al., 2017). children were excluded from the autistic group if they had a history of seizure disorders, progressive neurological disorders, hypoxic encephalopathy associated with prematurity, congenital rubella syndrome, neurofibromatosis, fetal alcohol syndrome, or cerebral palsy. Participant characteristics and details of group comparison can be found in subsequent chapters.

Procedure

The larger longitudinal study consisted of two time points approximately 12 months apart. Because data collection began during the COVID-19 pandemic, autism diagnostic, cognitive, and language standardized assessments were administered during a virtual (remote) visit. As public health restriction eased by Year 2, assessments were administered either remotely or in-person in the laboratory depending on clinician preference and/or health status. Because I was interested in anticipatory looking during a very narrow window in the prediction tasks of the larger project, I opted to administer the experimental eye-gaze tasks in-person in the laboratory rather than rely on virtual administration which had not yet been demonstrated to be reliable for use with clinical populations at the start of data collection.

Year 1

At Year 1, autistic children were between 2 and 3 years old while NT children were 1.5 to 3 years old. The experimental protocol at this time point consisted of one to two virtual visits for standardized assessment administration and two in-person visits for experimental eye-gaze tasks in addition to several caregiver report measures completed at home (described in Standardized Measures). At each of the two in-person visits, children participated in several eye-gaze experiments for the larger project in addition to those reported in the following studies. Of the studies comprising this dissertation, children only participated in the experiment reported in Chapter 3 at the first time point. The order of experimental eye-gaze task administration was counterbalanced across participants such that half the children saw the task the first day and half saw the task the second day. All children saw the task reported in Chapter 3 after the experimental tasks for the larger longitudinal project. The final visit(s) involved virtual administration of standardized assessments (described in Standardized Measures). Children in the NT group participated in one remote visit, while children in the autism group usually required two remote visits due to the length of the autism diagnostic measures. The NT remote visits lasted approximately two hours, while the autism group's remote visits lasted about 3.5 hours each. The in-person visits for the experimental eye-gaze tasks lasted approximately 45-60 minutes each for both groups. Parents received financial compensation according to the time required to complete the study protocol. Children in both groups also received a book for their participation.

Year 2

The second time point occurred approximately 12 months later, when autistic children were between 3 and 4 years old and NT children were between 2.5 and 4 years old. In addition to caregiver report measures completed at home, both groups participated in one visit conducted in-person in the laboratory at Visit 2. Clinicians sometimes opted for remote for standardized assessment administration depending on preference and health status. For both groups, this visit consisted of one experimental eye-gaze task for the larger project followed by standardized assessments and additional eye-gaze tasks, including those reported in the following studies. All experimental tasks reported in this dissertation were administered at Year 2. Once again, the order of experimental task administration order was counterbalanced across participants. In total, this visit lasted about 2-4 hours. Parents of the children enrolled in these studies received financial compensation based on the amount of participation time/sessions required. Children in both groups also received a book for their participation. Because some of the experimental tasks in this dissertation were only administered at Year 2, additional NT children in the Year 2 age range (2.5 to 4 years old) who did not participate in the larger parent project at Year 1 were recruited at that point to achieve an adequate sample size. These additional NT children participated in all experimental tasks reported in this dissertation, completed the same standardized assessments as those recruited for the larger study, and received the same compensation.

Standardized Measures

At Year 1, children completed a comprehensive developmental evaluation which included cognitive, language, adaptive behavior, and autism diagnostic measures. Caregivers also completed a background form to collect demographic information and developmental history. Due to the COVID-19 pandemic, our research team specifically selected measures that were valid and reliable for remote administration. In both groups, adaptive functioning was assessed through the Vineland Adaptive Behavior Scales, Third Edition (Vineland-3; Sparrow et al., 2016) Comprehensive Parent/Caregiver Form. Cognitive abilities were also assessed using a direct measure, the cognitive domain of the Developmental Assessment of Young Children – 2nd Edition (DAYC-2; Voress & Maddox, 2013). Children in both groups also received a direct measure of receptive and expressive language ability, the Preschool Language Scales – 5th Edition (PLS-5; Zimmerman et al., 2011). Children participated in a play-based language sample, though these data were not utilized in the present studies. The DAYC-2 and PLS-5 were all administered by a certified speech-language pathologist on our research team. To rule out symptoms of autism in the NT group, caregivers completed the Modified Checklist for Autism in Toddlers - Revised with Follow-Up (M-CHAT-R/F; Robins et al., 2014). To be included in the NT group, children had to receive passing scores (‘low risk’ scores of 0-2) on this screener. All children in the autistic group met criteria for DSM-5 autism diagnoses as determined by an experienced clinical psychologist on our research team. To determine diagnoses, the psychologist administered the Toddler Version of the Autism Diagnostic Interview – Revised (ADI-R; Kim & Lord, 2012), the Brief Observation of Symptoms of Autism (BOSA; Dow et al., 2022), and the Childhood Autism Rating Scale – Second Edition (CARS-2; Schopler et al., 2010). CARS-2 total scores were also used as a measure of level of autism traits in the analyses in Chapter 5.

At Year 2, caregivers of participants in both groups again completed the background form and the Vineland-3 prior to their laboratory visit. As in Year 1, children in both groups completed the DAYC-2 and the PLS-5 to measure cognitive and language abilities. In addition, children in the autistic group were administered the BOSA and CARS-2 at Year 2. At the time of data collection, there were no autism-specific screeners validated for children between 36-48 months and preliminary research into the adapted use of screeners for younger children had yielded poor specificity in this age group (Dolata et al., 2020; Salisbury et al., 2018). Therefore, autism symptoms in the additional NT children recruited at Year 2 who did not participate in Year 1 were ruled out based on parent report (i.e., no developmental delays, concerns or history of interventions on the background form), scores in the average range on standardized cognitive and language assessments, and clinical observation by a certified speech-language pathologist with autism experience.

Experimental Eye-Gaze Tasks

Procedure.

At both time points, children participated in experimental eye-gaze tasks conducted in-person in a sound-attenuated booth in our research lab. Before the experiment began, children were asked to watch and listen to the “movie.” Caregivers were instructed to avoid talking to their child, pointing, or otherwise directing their attention throughout the experiment. Caregivers wore opaque glasses to further prevent biasing their children’s gaze. Children sat on caregivers’ laps in a chair located approximately 60 cm from the screen in the booth. Visual stimuli appeared on the 55-inch screen while auditory stimuli played from a speaker under the screen at 65 dB. A Canon Vixia HFG10 video camera with a 30 Hz frame rate located just below the screen captured children’s eye movements. The video recording was monitored by the experimenter in real time from a Mac computer located just outside the booth. All experiments were programmed using E-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA) software on a PC computer.

Stimuli.

All eye-gaze tasks reported in this dissertation followed a Looking-While-Listening (LWL; Fernald, et al., 2008) design whereby children simultaneously heard auditory stimuli and viewed visual stimuli while their eye movements were recorded as an implicit measure of comprehension. Auditory stimuli were recorded by a female, native English speaker using child-directed intonation in the local regional dialect. The auditory stimuli consisted of single-syllable CVC novel words following English phonotactic constraints from the Novel Object and Unusual Name Database (NOUN Database; Horst & Hout, 2016) as well as a small number of familiar words appearing during the pre-training phase and twice during the test phase. Each task included four novel words and four novel objects. The audio was normalized such that the average intensity remained consistent across stimuli. During the training phase of each experiment, the novel words were presented either in pairs (in the cross-situational learning (CSL) tasks) or alone (in the ostensive teaching task) with no additional auditory cues. During the test phase of each experiment, the target word was spoken in the sentence frame (“*Where’s* *the* ____*?*") followed by a reinforcer sentence ("*Do you like/see it?*"). Visual stimuli consisted of photos of novel objects taken from the NOUN Database paired with the novel spoken words as well as photos of familiar objects corresponding to the familiar words. In half of the test trials, the target and distractor object photos were digitally altered in Adobe Photoshop (Adobe Inc., 2022) to change the hue of the image to a new color to assess generalization of learning (though these trials contained a design error and were not included in analyses). The objects appeared on 600 x 600 pixel grey boxes on a black background. In the CSL tasks, the boxes appeared side-by-side on screen during both the training and test phases. In the ostensive teaching task, a single box appeared in the middle of the screen during the training phase, while two boxes were side-by-side in the test phase as in the CSL tasks. For each experimental task, two versions were created such that trial order, target side, and label-image pairings were counterbalanced across children. Experiment structures were based on those used by Venker (2019) with some modifications. Trial orders can be found in Appendix B. Because the experimental design for the Chapter 4 study was within-subjects, each of the three tasks included a different set of novel words and images.

CSL Task Design.

The first study of this dissertation employed an experimental eye-gaze task following a very similar design to a prior study of cross-situational word learning (CSL) in autistic children (Venker, 2019). The task consisted of three phases: pre-training, training, and test. The purpose of the pre-training phase was to familiarize children with the structure of the task. In each of the four pre-training trials, two images of familiar objects (e.g., *car* and *book*) first appeared in grey boxes side-by-side on a black background in silence. After 500 ms, their labels were spoken with a 400 ms pause between the two labels. A 500 ms pause followed each trial. The familiar words used in this task were *car*, *book*, and *dog*. The training phase trials immediately followed the pre-training and followed the same structure and timing as the pre-training trials and included a total of four novel object images and four novel spoken labels, shown repeatedly throughout the training phase (Figure 1). The novel words taught in this task were *jick*, *tife, hux*, and *zeb*. The four novel objects used in this task are pictured in Figure 2. In each trial, two novel objects appeared side-by-side first in silence for 500 ms, then the novel labels were spoken with a 400 ms pause between them (e.g., "*hux… jick*"), with no auditory or visual cues linking the images and words. The audio files of the novel words were normalized to a duration of 1400 ms each. A 500 ms pause occurred between each trial. In total, each training trial lasted 4200 ms. The sequential order of the spoken labels and left-right orientation of the images were counterbalanced across training trials. The label-object mappings were intentionally ambiguous such that children must rely on co-occurrence statistics aggregated across trials to learn words. Each of the four novel word-object pairings appeared 10 times throughout the training phase over a total of 20 training trials. The training phase was immediately followed by the test phase. In each trial of the test phase, a target image and a distractor image (both images of novel objects which had appeared during the training phase) first appeared side-by-side in silence for 1000 ms. Then, the target word was spoken in the sentence frame "*Where's the ____?*" The onset of the target word occurred 1000 ms after the beginning of the sentence frame audio across test trials. This was followed by a 450 ms pause, then a reinforcer sentence "*Do you see/like it?*" that lasted 1500 ms. In total, each test trial lasted 5650 ms. Each novel word was presented as the target four times during the test phase. Two of the test trials included the same image of the novel object that had appeared during training to assess fast mapping, while the other two included a version of the novel object digitally altered to a different color to assess generalization of learning. Test trials were counterbalanced for target side and presentation order. It should be noted that in the initial version of the experiment, the generalization test trials paired the untrained (digitally altered) target image with a distractor that had appeared during training. This error in design therefore confounded looking due to learning with looking due to image novelty. Because this error was discovered after most of the data had been collected, I was not able to analyze the generalization data as intended. Two additional test trials included familiar objects to introduce the trial structure and support children's attention to the task for a total of 18 test trials. Every 6 trials, a short (5 sec) video played to maintain children's attention. In total, the task lasted approximately 3.5 minutes.

CSL with Exemplar Variability Task Design.

The task reported in Chapter 4 investigating the effect of exemplar variability in autistic children's word learning included two additional experimental eye-gaze tasks. The first task followed the same structure as the CSL task described above with one crucial difference: throughout training, each novel word was paired with three different images of novel objects that differed in color but shared the same shape (Figure 3). Because the sets of images in Multiple Exemplars section of the NOUN Database were limited and already used in several other eye-gaze tasks in the larger project administered contemporaneously to these studies, two of the four novel words in each of the CSL+V and O+V tasks (described below) were paired with a set of three images from the Multiple Exemplars section of the NOUN Database (Figure 4), and the other two were paired with a set of three images originating from a single novel object image that were digitally altered in Adobe Photoshop to change their color scheme. The familiar words used in this task were *ball*, *cup*, and *shoe*, and the novel words used in this task were *poss*, *dax*, *mel*, and *gip*. The CSL with exemplar variability (hereafter CSL+V) task was identical to the basic CSL task described above in timing, number/types of trials, counterbalancing, number of novel words, and CVC single-syllable word structure. As such, I was able to isolate the effect of variability in the exemplar images during training when comparing learning during test trials between the two CSL tasks. This task contained the same experimental design error as the first study, and as such generalization test trials were not analyzed.

Ostensive Teaching with Exemplar Variability Task Design.

A second experimental eye-gaze task was administered to determine whether children learn more easily from exemplar variability in an ostensive (unambiguous) learning context compared to the CSL+V task. This task mirrored the CSL+V task with one critical design difference: in each trial during the pre-training and training phases, children only saw one object on the screen and heard one label spoken at a time. To ensure the total number of exposures to each novel word was identical between the CSL+V and ostensive teaching with exemplar variability (hereafter O+V) tasks, there were 40 training trials in the O+V task (10 per word). Like the CSL+V task, the training trials began with the image on screen. After 500 ms of silence, the audio of the spoken novel word played, lasting 1400 ms, followed by a 500 ms blank screen before the next trial. The test trials were identical in overall design and timing to the CSL+V task, though trial order differed. This task also shared the experimental design error noted in the other tasks, and as such I was not able to analyze generalization trials. A 5-second video played after every 6 trials to help maintain children's attention to the task. The total duration of the O+V task was the same as the CSL+V tasks (approximately 3.5 minutes). Two different versions of the task were created such that trial order, target image side, and novel word-object pairings were counterbalanced. The familiar words used in the O+V task were *ball*, *cup*, and *shoe*, and the novel words used were *dage*, *juff*, *lep*, and *shede*. Three images were used throughout the training for each novel word. The three images given the same novel label all shared the same shape but differed in color. Novel visual stimuli appearing during training for all experimental eye-gaze tasks can be found in Appendix A, and the two trial orders for each of the three eye-gaze tasks can be found in Appendix B.

Data Analysis

Data Processing

Video recordings of children's eye movements during the experimental tasks were coded offline by trained research assistants in peyeCoder (Olson, et al., 2020). Coders were not aware of visual or auditory stimuli while coding and followed established guidelines for manual eye-gaze coding (Fernald et al., 2008). As they viewed the video footage of children's gaze behavior, the research assistants assigned each 33-ms video frame a code depending on whether the child was looking at the right image, left image (or center image during the training trials of the O+V only), was shifting their gaze, or was looking away. A second coder independently re-coded a randomly-selected portion of the videos. Inter-coder reliability was defined by the extent to which the second coder agreed with the original coder through two measures: frame agreement (the average proportion of frames in which coders agreed) and shift agreement (the average proportion of frames the coders agreed the child was shifting their gaze). Across all children tested in the CSL task at Year 1, frame agreement was 96.28% for the autistic group and 94.45% for the NT group while shift agreement was 93.46% for the autistic group and 90.86% for the NT group. Across all children tested and coded in the CSL task at Year 2, frame agreement was 93.16% for the autistic group and 93.70% for the NT group while shift agreement was 89.14% for the autistic group and 91.26% for the NT group. Across all children tested and coded in the CSL+V task (administered at only Year 2), frame agreement was 86.79% for the autistic group and 94.93% for the NT group while shift agreement was 87.35% for the autistic group and 91.71% for the NT group. Finally, frame agreement was 93.95% for the autistic group and 96.64% for the NT group across all children tested and coded in the O+V task (administered at only Year 2) while shift agreement was 87.85% for the autistic group and 91.08% for the NT group. Note that data were collected for four additional NT participants at Year 2 but were not analyzed due to time constraints.

Data Cleaning

Following the same data cleaning procedure of previous eye-gaze studies (e.g., Venker, 2019; Prescott et al., 2022), test trials were excluded from analyses if children were looking to the visual stimuli on screen for less than 50% of analysis window (300-1800 ms after target noun onset). For each task, children were excluded from analyses if they contributed 2 or fewer (of 8 total) trained (i.e., fast mapping) test trials. Average number of trials contributed by participants in each group varied by task and can be found in following chapters.

Analytic Approach

Statistical analyses were performed using R (version 4.3.1; R Core Team, 2023) and RStudio (version 2023.6.2.561; Posit Team, 2023) software. Across studies, I focused my analyses on a window of 300 ms-1800 ms after target noun onset in the test trials, in line with many Looking-While-Listening studies (Fernald et al., 2008). In doing so, children were allowed sufficient time to generate a visual saccade in response to the target noun (Canfield et al., 1997). Details of data analysis varied by study and can be found in subsequent chapters.

References

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K., & Pickles, A. (2007). Patterns of Growth in Verbal Abilities Among Children With Autism Spectrum Disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. https://doi.org/10.1037/0022-006X.75.4.594

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., Snow, K. L., Aslin, R. N., Haith, M. M., Wass, T. S., & Adler, S. A. (1997). Information Processing Through the First Year of Life: A Longitudinal Study Using the Visual Expectation Paradigm. *Monographs of the Society for Research in Child Development*, *62*(2), 1–160.

Dolata, J. K., Sanford-Keller, H., & Squires, J. (2020). Modifying a general social-emotional measure for early autism screening. *International Journal of Developmental Disabilities*, *66*(4), 296–303. https://doi.org/10.1080/20473869.2019.1577024

Dow, D., Holbrook, A., Toolan, C., McDonald, N., Sterrett, K., Rosen, N., Kim, S. H., & Lord, C. (2022). The Brief Observation of Symptoms of Autism (BOSA): Development of a New Adapted Assessment Measure for Remote Telehealth Administration Through COVID-19 and Beyond. *Journal of Autism and Developmental Disorders*, *52*(12), 5383–5394. https://doi.org/10.1007/s10803-021-05395-w

E-Prime 3.0. (2016). [Computer software]. Psychology Software Tools, Inc. https://support.pstnet.com/

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. M. Fernández, & H. Clahsen (Eds.), *Developmental*

Psycholonguistics: On-line methods in children's language processing (pp. 97–135). John Benjamins. https://www.researchgate.net/publication/267179424

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.

Kim, S. H., & Lord, C. (2012). New autism diagnostic interview-revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders*, *42*(1), 82–93. https://doi.org/10.1007/s10803-011-1213-1

Olson, R. H., Pomper, R., Potter, C. E., Hay, J. F., Saffran, J. R., Ellis Weismer, S., & Lew-Williams, C. (2020). *Peyecoder: An open-source program for coding eye movements* (v1.0.2-beta) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.3939234

Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., & Fein, D. (2014). *Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F)*. *133*(1).

Salisbury, L. A., Nyce, J. D., Hannum, C. D., Sheldrick, R. C., & Perrin, E. C. (2018). Sensitivity and Specificity of 2 Autism Screeners Among Referred Children Between 16 and 48 Months of Age. *Journal of Developmental & Behavioral Pediatrics*, *39*(3), 254. https://doi.org/10.1097/DBP.0000000000000537

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale—Second Edition (CARS-2)*. Western Psychological Services.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3)*. Pearson.

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*(November 2018), 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Voress, J. K., & Maddox, T. (2013). *Developmental Assessment of Young Children-Second Edition (DAYC-2)*. PRO-ED.

Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., & Naigles, L. R. (2017). Grammatical Language Impairment in Autism Spectrum Disorder: Exploring Language Phenotypes Beyond Standardized Testing. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00532

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Tables and Figures

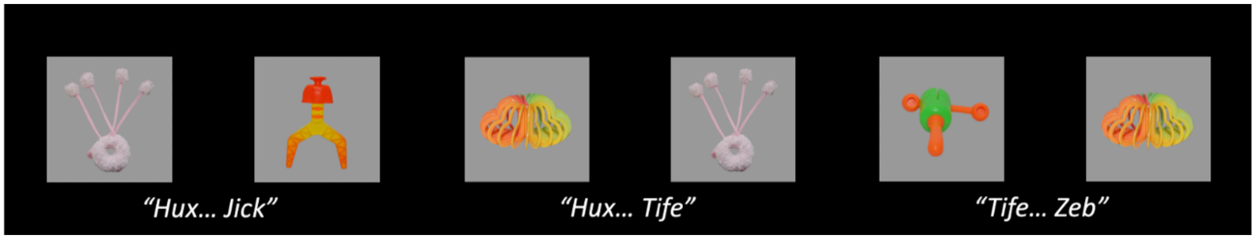


Figure 1. Example of three CSL task trials. Auditory stimuli are represented by text in quotations.

Figure 2. Visual stimuli used during the training phase in the CSL task.

Figure 4. Example of three images all given the same novel label (e.g., *dax*) during the training phase of the CSL+V and O+V experimental eye-gaze tasks.

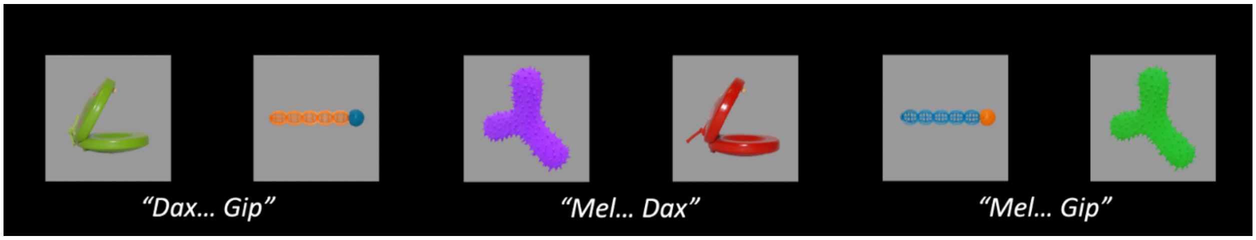


Figure 3. Example of three CSL+V task trials. Auditory stimuli are represented by text in quotations.

Chapter 3: Cross-Situational Word Learning in Young Autistic Children

The first study of this dissertation investigated cross-situational word learning (CSL) in 2- to 4-year-old autistic children and NT peers. Using an eye-tracking paradigm, I sought to examine novel word learning acquired through a classic CSL paradigm as described in General Methods (Chapter 2) in young autistic children and compare their performance to NT children. I asked the following research questions:

1. Are 2- to 4-year-old autistic children able to learn novel words through CSL, as measured by accuracy (i.e., proportion of looking to target images) greater than chance (.50)?
2. Do 2- to 4-year-old autistic children differ from NT peers in their novel word learning acquired through CSL?

Methods

Participants

As described in General Methods (Chapter 2), most children were recruited for this dissertation from an ongoing parent project in the lab. Across all participants tested in this task who passed exclusionary criteria, 11 autistic children and 14 NT children were tested at the Year 1 time point of the larger longitudinal study (when autistic children were 2- to 3-years-old) while 20 autistic children and 12 NT children were tested at the Year 2 time point (when autistic children were 3- to 4-years-old). An additional 17 NT children who were in the age range of the Year 2 time point (but did not participate in the parent project at Year 1) were recruited to participate in just dissertation tasks, resulting in a total of 29 NT participants at Year 2. 6 NT children and 2 autistic children were tested in this task at both Year 1 and Year 2, but only Year 2 data were included for these participants to avoid non-independence in the data. One additional

NT participant participated in other tasks for this dissertation but was too fussy to attempt the current task. Four additional NT participants were tested in the dissertation tasks as part of the visit protocol for the larger parent study at the Year 2 time point but were not included as data analysis for this dissertation had already begun. As part of the larger study plan, NT children were recruited intentionally younger to facilitate matching on nonverbal cognitive ability.

Following data cleaning (described below), remaining autistic and NT children in the sample were matched on Developmental Assessment of Young Children – Second Edition (DAYC-2; Smith & Yu, 2008) cognitive domain raw scores using a 1:1 bootstrap procedure previously used by our research group in other studies (Pomper et al., 2019; Prescott et al., 2022). This procedure used the bootmatch package in R (https://github.com/tjmahr/bootmatch). The procedure ran 100 bootstrap samples with a caliper of 5 (that is, children in each matched pair could differ by no more than 5 points on the DAYC-2 cognitive domain raw). I opted to match on nonverbal cognitive ability rather than a language measure for two reasons. First, the majority of the participants were originally recruited as part of a larger study with a planned design to match on nonverbal cognitive ability. As a result, the range of cognitive ability in the NT group was a reasonable match for the range in the autistic group, whereas the range of absolute language ability was considerably higher (see Table 1 and 2). Therefore, attempting to match groups on Preschool Language Scales 5th Edition (PLS-5; Zimmerman et al., 2011) Auditory Comprehension scores would result in inadequate sample sizes. Second, many researchers have argued that statistical learning is at least partially supported by domain-general mechanisms (Frost et al., 2015; Kirkham et al., 2002; Krogh et al., 2013) and CSL in particular is related to multiple cognitive domains in NT children (Vlach & DeBrock, 2017). This procedure resulted in *n*=19 in the autistic group (5 female), *m*=40.68 months, *SD*=7.09, range=27-48) and *n*=19 in the NT group (7 female), m=27.63 months, *SD*=6.45, range=19-39). While significantly different on age (*p*<.001), these groups did not differ on nonverbal cognitive ability (*p*=.266). See Table 1 for full participant characteristics of the cognitive ability-matched sample.

Because neither nonverbal cognitive ability-matched group demonstrated learning as measured by a group average proportion looks to target significantly greater than .50 in this task (see results below), I then conducted a secondary analysis comparing all children tested in this task at the Year 2 time point under the assumption that older children would generally perform better on this task than younger children as was found in previous CSL studies with NT children (Akhtar & Montague, 1999; Vlach & DeBrock, 2019). This analysis included all 17 autistic children (5 female; *m*=44.00 months, *SD*=4.17, range=36-50) and 27 NT children (13 female, *m*=32.96 months, *SD*=2.26, range=30-39) who participated in the task at the Year 2 time point, passed exclusionary criteria, and contributed sufficient data for inclusion (see Data Cleaning). While these groups significantly differed on age, nonverbal cognitive ability, and receptive language ability, I included nonverbal cognitive ability as a covariate as described in Analytic Approach. Child characteristics and demographics for each group at Year 2 can be found in Table 2.

Data Analysis

Data Cleaning

I removed all 33-ms video frames in which children were shifting their gaze or were looking away from the visual stimuli from analyses such that the final accuracy measure represented the average proportion of looks to target images over looks to target and distractor images during the analysis window (300-1800 ms after target noun onset in test trials, which occurred 2000 ms into the trial (Figure 5). As in several previous LWL studies (e.g., (Mathée-

Scott et al., 2022; Pomper et al., 2019; Venker et al., 2021) I excluded test trials in which children were looking to visual stimuli for less than 50% of the analysis window. I then excluded children who contributed 2 or fewer out of a total 8 trained (i.e., fast mapping trials; 25%) to reduce possible bias in the average accuracy measure (e.g., a child who contributed only one or two trials may have been very accurate in those trials but uncertain in other trials and therefore avoided looking at the visual stimuli, resulting in exclusion of those trials from analyses after data cleaning and artificially inflating their average accuracy aggregated across trials). Note that untrained (i.e., generalization) test trials were not usable in analyses due to an experimental design error. Therefore, while the task included 2 familiar test trials (to introduce trial structure and retain children's attention), 8 untrained (i.e., generalization) test trials and 8 trained (i.e., fast mapping) test trials, I only proceeded with analyses using the 8 trained (fast mapping) trials. Across all children tested in the task at the Year 1 time point after data cleaning and exclusion, autistic children contributed an average of 7/8 test trials (*SD*=1, range=6-8) and NT children contributed an average of 7/8 test trials (*SD*=1, range=6-8). Across all children tested at the Year 2 time point after data cleaning and exclusion, autistic children contributed an average of 5/8 test trials (*SD*=1, range=3-8) and NT children contributed an average of 7/8 test trials (*SD*=1, range=4-8).

Analytic Approach

After cleaning, eye-gaze data from within the analysis window (300-1800 ms after target noun onset) across all trained (i.e., fast mapping) test trials were averaged within each participant, resulting in one measure of accuracy (i.e., proportion looks to target) per participant. Raw looking data over time during the analysis window is depicted in the profile plots for the matched groups (Figure 6) and the children included in the secondary analysis (Figure 7) but the aggregated accuracy data used in analyses is pictured in the boxplots (Figure 8 and 9). For the first research question examining whether children in each group were able to learn novel words using cross-situational statistics, I followed the analytic approach used by Venker (2019) and conducted one-sample, two-tailed t-tests comparing each group's average accuracy against chance (.50). For the second research question investigating group differences on task performance, I used a two-sample, two-tailed t-test for the initial analysis with nonverbal cognitive ability-matched groups as in Venker (2019). However, because groups were not matched in the secondary analysis of Year 2 children, I fit a regression model including nonverbal cognitive ability as a covariate. Standard ordinary least squares (OLS) multiple linear regression is not generally recommended for proportional data like the accuracy variable because it is bounded by 0 and 1 (Cribari-Neto & Zeileis, 2010). Therefore, to accommodate the bounded nature of the outcome variable, I fit a generalized linear model using a beta distribution and logit link with the *glmmTMB* package in R (Brooks et al., 2017) which regressed average accuracy for each child (in terms of log-odds of looking to target due to logit link) on diagnostic group (contrasts: NT=-0.5, autistic=0.5) and nonverbal cognitive ability (as measured by DAYC-2 cognitive domain raw scores, grand-mean scaled and centered) as a covariate to account for group differences on this variable.

Results

Nonverbal Cognitive Ability-Matched Groups

One-sample two-tailed t-tests revealed that neither the autistic group (m=0.56, $t(18)$=1.35, p=.19) nor the NT group (m=0.54, $t(18)$=1.10, p=.28) looked to target images significantly above chance (0.50) in the matched sample. The groups also did not significantly differ in average proportion looks to target, ($t(34.25)$=-0.37, p=.71).

Full Sample at Year 2 Time Point

Among autistic children tested at Year 2 who passed data cleaning and exclusionary criteria, proportion target looks did not exceed chance, (*m*=0.54, *t*(16)=0.83, *p*=.42). Among NT children tested at Year 2 who passed data cleaning and exclusionary criteria, proportion looks to target was significantly above chance, (*m*=0.58, *t*(26)=3.08, *p*=.005). However, model results indicated that when nonverbal cognitive ability was included as a covariate, groups did not significantly differ in task performance, (*β*=0.03, *SE*=0.27, *z*=0.13, *p*=.900). Full model results can be found in Table 3.

Discussion

The present study sought to examine CSL in preschool-age autistic children and compare their learning to NT peers. Children participated in an eye-gaze experiment where they heard two novel words while viewing two novel objects with no cues linking the words with their correct referents. Therefore, children had to track the statistical co-occurrences of words and objects across contexts (i.e., trials) to learn their correct mappings. In test trials, children saw the target object and a distractor object side-by-side on screen while they were asked *"Where's the **hux**? Do you like/see it?"* while their eye movements were video recorded and coded offline. Word learning was operationalized as proportion of looks to target over looks to target and distractor within an analysis window 300-1800 ms after target noun onset averaged across test trials for each child (i.e., accuracy).

Unlike previous studies with older children (e.g., Hartley et al., 2020; Venker, 2019), this study did not find evidence of CSL in autistic children as a group during the preschool years, at least within this paradigm as designed. Importantly, younger NT children matched to the autistic group on nonverbal cognitive ability also did not show evidence of CSL as a group, and matched groups did not differ in their accuracy. This finding suggests that while autistic children may not have learned via CSL as defined by this study design, NT children of a similar cognitive ability level did not show superior performance. Because I assumed older children would perform better than younger children in this task similar to previous research with NT children (Akhtar & Montague, 1999; Vlach & DeBrock, 2019; Vlach & Johnson, 2013), I opted to conduct a secondary analysis comparing groups of children tested at the Year 2 time point. All children tested at Year 2 who passed data cleaning and exclusionary criteria were included in this analysis. One-sample t-test results indicated that NT children at this time point were able to learn words via CSL as measured by average accuracy significantly above chance, while the autistic children at Year 2 did not show above-chance accuracy. However, it should be noted that these groups were not matched on any variable, and in fact significantly differed on every child characteristic measured in this study (Table 2). When directly compared and including nonverbal cognitive ability as a covariate, groups did not differ on average accuracy.

While I had anticipated that autistic children in this age range may show relatively more difficulty learning words from CSL, results may suggest that this particular paradigm was difficult for children in the early stages of development such that even NT children at an average age of 27.63 months (range 19-39; i.e., the cognitive ability-matched sample) did not show evidence of learning as a group on this task. The Year 2 sample of NT children, who did show above-chance accuracy as a group, were 32.96 months on average (range 30-39). This pattern of findings was surprising as previous studies have reported that NT children as young as 12 months were able to learn words via CSL (Smith & Yu, 2008; Vlach & DeBrock, 2017; Vlach & Johnson, 2013). Methodological and analytical differences might account for these discrepant findings. In Smith and Yu (2008), test trials were 8 seconds long, and target words were repeated four times. While there were even more target noun-referent pairs than the present study (6 vs. 4), the visual stimuli were simpler, consisting of different colored two-dimensional shapes. The outcome variable was duration of looks to target and distractor, and the analytic approach was a mixed analysis of variance. Any of these differences may have boosted the measure of learning in Smith and Yu (2008) compared to the present study. In particular, the longer test trials with multiple target word repetitions afforded children more opportunities to direct their gaze to target, whereas the present study examined a window only 1500 ms long following a single production of the target word. Vlach and Johnson (2013) reported that 16- and 20-month-old NT children were able to learn novel words at above-chance levels in a CSL task, though their methods also differed from the present study. They compared massed and interleaved learning phase schedules presented in a block design, with test trials that were 8 seconds long. Vlach and DeBrock (2017) found evidence of CSL in 2- to 3-year-old and 4- to 5-year-old NT children, but again differed considerably in study design. For one, children were required to point to the target image on test trials, and accuracy was then measured as percent correct responses. This resulted in an attrition rate of 32% due to fussiness, inability to follow directions, etc. (Vlach & DeBrock, 2017). The final included sample in Vlach and DeBrock (2017) was therefore limited to children who were able to participate in a behaviorally demanding task, which likely contributed to the higher rate of accuracy reported compared to the present study, which did not require any explicit behavioral response. Venker (2019) provides a closer comparison to the current study in terms of experimental design. The NT sample in Venker (2019) was 58 months on average but included a range from 31-95 months. The present study confirms that NT children at the very youngest end of this range were also able to successfully learn words via CSL in this paradigm. Meanwhile, the autistic sample in Venker (2019) was 76 months on average (range=48-95).

Thus, while the youngest children in Venker's (2019) autistic group were at the maximum of the range of the present study, they were on average considerably older. While null results should be interpreted cautiously, it is possible that successful learning in this CSL paradigm comes online for autistic children at an age between that of the present sample and Venker (2019). Alternatively, children in this age range might have demonstrated learning in a different task design, perhaps given more repetitions and/or time to shift gaze as in Smith and Yu (2008). It is worth noting the considerable variability in task performance among the autistic children; visual inspection indicates that even in this relatively rigorous task, several children were individually well above chance in their average proportion looks to target (Figure 8; Figure 9). Predictors of success in this task for the autistic group will be further explored in Chapter 5.

Aside from experimental design and developmental level, other factors may have contributed to the null findings of learning by autistic children in this study. First, autistic children may not have attended to the training trials to the same extent as NT children. To investigate this possibility further, I defined lack of attention to (i.e., missing) training trials as average percent of frames in which children were shifting their gaze or looking away from visual stimuli during a window from the onset of the first spoken word to the end of the second word across training trials. Among all children tested on the task who passed exclusionary criteria (before data cleaning), NT children tested at Year 1 on average missed 14.15% of training trials (*SD*=8.10, range=3.71-23.04) while NT children tested at Year 2 missed 20.41% (*SD*=15.17, range=5.62-65.52). By contrast, autistic children tested at Year 1 missed 22.37% (*SD*=14.33, range=6.9-44.28) while autistic children tested at Year 2 missed 34.52% (*SD*=17.52, range=8.66-71.44). Autistic children have been noted to demonstrate differences in visual attention (Elsabbagh et al., 2013; Sacrey et al., 2013) which could lead some children to make different or weaker word-referent mappings. The profile plot of Year 2 children's proportion target looks over time (Figure 7) may also hint at differences in visual attention between groups. Describing qualitatively, the plot suggests the autistic group may have begun looking to the target earlier than NT children, even before the start of the analysis window. However, they cease looking more to target about halfway through the analysis window, while the NT children continue to look more to the target for its duration. Thus it is possible that autistic children did in fact learn, but that learning was not captured by this operationalization due to differences in the timing of their gaze behavior. A final possible explanation for null findings in this study was the relatively demanding visit protocol of the larger longitudinal study in which these data were collected. At Year 1, children always saw another experimental video before the present study. At Year 2, most participants saw up to three experimental videos before the present study, several of which were also word learning experiments. It is conceivable that fatigue negatively impacted children's accuracy and ability to attend to this task as well.

Limitations and Future Directions

Several limitations should be noted for the present study which could be addressed in future work. First, this sample was predominantly white and non-Hispanic or Latino, and the NT sample was significantly higher than autistic peers in maternal years of education (often considered a proxy for socioeconomic status). This discrepancy is expected due to differences in recruiting between groups. Across all children tested in this task who met inclusionary criteria, all but one autistic child was recruited from across the broader region outside of Dane County, WI, while all but one NT children was recruited from within Dane County, WI. Therefore, the findings of this study may not be as generalizable and should be investigated further in future studies with a more representative sample. By contrast, a strength of this study was the inclusion of autistic children with a wide range of ability, better representing the heterogeneity of the autism spectrum. However, this choice resulted in considerable differences in child characteristics between the autistic and NT groups, even in the cognitive ability-matched sample (Table 1 and Table 2). Comparing this more representative autism sample to NT peers posed challenges. Even when groups were matched on nonverbal cognitive ability, they differed in age, expressive and receptive language ability, and adaptive behavior. The NT group who did show evidence of learning (i.e., those tested at Year 2 who passed data cleaning and exclusionary criteria) significantly differed from the autistic children tested at Year 2 across all child characteristics reported (Table 2). Attempts to match the children tested at Year 2 on nonverbal cognitive ability were unsuccessful, resulting in too few children per group. Instead, I included nonverbal cognitive ability as a covariate in the analysis. It should be noted that both matching groups and covarying have been questioned as appropriate solutions to the problem of differences in diagnostic group characteristics such as cognitive or language ability (Dennis et al., 2009; Miller & Chapman, 2001; Plante et al., 1993). For example, if the groups are matched on absolute cognitive or language ability, at least some children in the NT group must be younger than the autistic group. If the groups are the same age but the autistic group includes children with the full range of cognitive or language ability, the groups will differ on those characteristics. Covarying child characteristics is an alternative option, but researchers have argued this approach may remove variance that is meaningfully and inextricably tied to diagnostic group and produce uninterpretable results (Miller & Chapman, 2001). In this study, I conducted two analyses using both approaches. Matching on cognitive ability resulted in a NT comparison group that was too young to successfully learn in the task. The secondary analysis covaried cognitive ability with a NT group that was able to successfully learn in the task and an autistic group who were not. This analysis did not yield group differences. Future studies examining CSL in autistic children should carefully consider the characteristics on which they would like to compare groups and ensure that the experimental design yields learning in NT children in that developmental range before proceeding. Another limitation to note is the possible bias introduced by my decision to exclude participants who contributed two or fewer usable trials. While this decision avoided some bias by preventing inclusion of less stable (i.e., based on fewer data) measures of accuracy, the missingness of these data may not be random if children who more consistently looked away from the visual stimuli differed in meaningful ways from those who looked more to visual stimuli. A final limitation of this study was the experimental design error on generalization test trials which necessitated their removal from analyses. For all but five NT children who saw a corrected version of the task, generalization test trials included a target image that was a novel color paired with a distractor image in the same color it had originally appeared during training. As a result, preliminary visual inspection of the data indicated that children tended to look very reliably to the target image, likely simply because it was a novel color. An elegant experimental design assessing generalization of word learning from cross-situational statistics can be found in Hartley et al. (2020). In that study, generalization test trials consisted of a target object “shape match” in a new color and a “color match” distractor object that was the same color as the trained target image but the shape of one of the other trained images. Future studies might consider using a similar design as Hartley et al. (2020) with autistic children in the younger age range of the present study.

Conclusions

This study did not find evidence of CSL in 2- to 4-year-old autistic children as a group, at least in this paradigm. However, younger, nonverbal cognitive ability-matched NT children also did not demonstrate learning and did not differ from autistic peers. NT children who were slightly older on average (*m*=32.96 months) did show evidence of learning but were difficult to compare to autistic peers tested at the same time point due to large group differences in cognitive and language abilities. Individually, autistic children varied considerably in task performance (Figure 8 and Figure 9) consistent with the heterogeneity in many abilities across the autism spectrum. Predictors of individual differences in CSL among autistic children will be explored further in Chapter 5.

References

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347–358. https://doi.org/10.1177/014272379901905703

Brooks, M., E., Kristensen, K., Benthem, K., J. ,van, Magnusson, A., Berg, C., W., Nielsen, A., Skaug, H., J., Mächler, M., & Bolker, B., M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378. https://doi.org/10.32614/RJ-2017-066

Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in *R*. *Journal of Statistical Software*, *34*(2). https://doi.org/10.18637/jss.v034.i02

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*(3), 331–343. https://doi.org/10.1017/S1355617709090481

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of Visual Attention in Infancy is Associated with Emerging Autism in Toddlerhood. *Biological Psychiatry*, *74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*(July 2019), 104265. https://doi.org/10.1016/j.cognition.2020.104265

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Krogh, L., Vlach, H., & Johnson, S. (2013). Statistical Learning Across Development: Flexible Yet Constrained. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00598

Mathée-Scott, J., Larson, C., Venker, C., Pomper, R., Edwards, J., Saffran, J., & Ellis Weismer, S. (2022). Use of Mutual Exclusivity and its Relationship to Language Ability in Toddlers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *52*(10), 4528–4539. https://doi.org/10.1007/s10803-021-05321-0

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48. https://doi.org/10.1037/0021-843X.110.1.40

Plante, E., Swisher, L., Kiernan, B., & Restrepo, M. A. (1993). Language matches: Illuminating or confounding? *Journal of Speech and Hearing Research*, *36*(4), 772–776. https://doi.org/10.1044/jshr.3604.772

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

Prescott, K. E., Mathée-Scott, J., Reuter, T., Edwards, J., Saffran, J., & Ellis Weismer, S. (2022). Predictive language processing in young autistic children. *Autism Research*, *15*(5), 892–903. https://doi.org/10.1002/aur.2684

Sacrey, L. A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural Brain Research*, *256*, 441–450. https://doi.org/10.1016/j.bbr.2013.08.028

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale—Second Edition (CARS-2)*. Western Psychological Services.

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3/4), 591–611.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3)*. Pearson.

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*(November 2018), 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Venker, C. E., Mathée, J., Neumann, D., Edwards, J., Saffran, J., & Ellis Weismer, S. (2021). Competing Perceptual Salience in a Visual Word Recognition Task Differentially Affects Children With and Without Autism Spectrum Disorder. *Autism Research*, *14*(6), 1147–1162. https://doi.org/10.1002/aur.2457

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230. https://doi.org/10.1016/j.jml.2016.10.001

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 700–711. https://doi.org/10.1037/xlm0000611.Statistics

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015.Memory

Voress, J. K., & Maddox, T. (2013). *Developmental Assessment of Young Children-Second Edition (DAYC-2)*. PRO-ED.

Wilcoxon, F. (1945). Some Uses of Statistics in Plant Pathology. *Biometrics Bulletin*, *1*(4), 41–45. https://doi.org/10.2307/3002011

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Tables and Figures

Table 1. Participant characteristics and demographic information for cognitive ability-matched sample.

	Autistic Group (n=19)	NT Group (n=19)	Group Comparison
	Mean (SD) Median Range	Mean (SD) Median Range	
Participant Characteristics			
Age (Months)	40.68 (7.09) 45 27-50	27.63 (6.45) 31 19-39	*p*<.001
Receptive Language			
Raw Score	24.47 (10.67) 20 11-53	36.82 (13.01) 33 19-65	*p*=.001
Standard Score	64.95 (20.17) 57 50-125	114.88 (15.28) 117 81-139	*p*<.001
Expressive Language			
Raw Score	27.32 (7.85) 24 16-44	33.29 (7.49) 36 24-49	*p*=.03
Standard Score	73.89 (15.71) 71 52-103	111.00 (11.40) 106 98-130	*p*<.001
Cognitive Ability			
Raw Score	37.63 (7.87) 35 28-57	39.74 (7.26) 40 30-60	*p*=.266
Standard Score	81.37 (14.01) 80 57-116	104.37 (7.24) 103 96-120	*p*<.001
Adaptive Behavior	70.42 (9.71) 69 53-97	99.21 (9.95) 96 80-121	*p*<.001

Level of Autism Traits	34.61 (4.51) 34.5 26.5-43.5	—	—
Demographic Information			
Maternal Education (Years)	13.53 (2.20) 12 11-18	17.53 (2.37) 17 14-24	$p<.001$
Race	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 1 Black or African American 5 More than one race 13 White	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 0 More than one race 19 White	
Ethnicity	2 Hispanic or Latino 17 Not Hispanic or Latino	0 Hispanic or Latino 19 Not Hispanic or Latino	

Note. Group differences were quantified by two-sample, two-tailed t-tests if both groups were normally distributed as measure by Shapiro-Wilk normality test (Shapiro & Wilk, 1965) or Wilcoxon rank sum tests (Wilcoxon, 1945) with continuity correction if either was not. Resulting p-values are reported in the third column. Receptive language was measured by PLS-5 Auditory Comprehension subtest scores (Zimmerman et al., 2011). Expressive language was measured by PLS-5 Expressive Communication subtest scores. Two NT participants did not complete the PLS-5. Cognitive ability was measured by DAYC-2 Cognitive Domain scores. Adaptive Behavior was measured by Vineland-III Adaptive Behavior Composite Scores (Sparrow et al., 2016). Level of autism traits was measured by CARS-2 total scores, where 15-29.5 indicates minimal-no autism traits, 30-36.5 indicates mild-moderate autism traits, and 37 and higher indicates high autism traits (Schopler et al., 2010).

Table 2. Participant characteristics and demographic information for all participants at the Year 2 time point included in secondary analysis.

	Autistic Group (n=17)	NT Group (n=27)	Group Comparison
	Mean (SD) Median Range	Mean (SD) Median Range	
Participant Characteristics			
Age (Months)	44.00 (4.17) 46 36-50	32.96 (2.26) 33 30-39	*p*<.001
Receptive Language			
Raw Score	24.76 (11.08) 20 11-53	43.29 (7.68) 42 33-65	*p*<.001
Standard Score	63.47 (21.35) 53 50-125	118.71 (12.11) 119 95-142	*p*<.001
Expressive Language			
Raw Score	26.94 (8.30) 24 13-44	38.96 (4.08) 38 33-49	*p*<.001
Standard Score	70.18 (15.72) 63 50-103	112.38 (9.02) 111 97-129	*p*<.001
Cognitive Ability			
Raw Score	37.24 (8.04) 35 28-57	45.48 (4.38) 44 40-60	*p*<.001
Standard Score	75.35 (12.24) 74 56-104	107.26 (7.05) 105 98-122	*p*<.001
Adaptive Behavior	67.12 (7.87) 68 53-79	103.33 (12.17) 103 80-131	*p*<.001

Level of Autism Traits	35.91 (4.81) 35 26.5-47	—	—
Demographic Information			
Maternal Education (Years)	13.59 (2.32) 12 11-18	17.30 (2.48) 16 14-26	$p<.001$
Race	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 0 More than one race 27 White	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 6 More than one race 11 White	
Ethnicity	0 Hispanic or Latino 27 Not Hispanic or Latino	2 Hispanic or Latino 15 Not Hispanic or Latino	

Note. Group differences were quantified by two-sample, two-tailed t-tests if both groups were normally distributed as measure by Shapiro-Wilk normality test (Shapiro & Wilk, 1965) or Wilcoxon rank sum tests (Wilcoxon, 1945) with continuity correction if either was not. Resulting p-values are reported in the third column. Receptive language was measured by PLS-5 Auditory Comprehension subtest scores (Zimmerman et al., 2011). Expressive language was measured by PLS-5 Expressive Communication subtest scores. Three NT participants did not complete the PLS-5. Cognitive ability was measured by DAYC-2 Cognitive Domain scores. Adaptive Behavior was measured by Vineland-III Adaptive Behavior Composite Scores (Sparrow et al., 2016). Level of autism traits was measured by CARS-2 total scores, where 15-29.5 indicates minimal-no autism traits, 30-36.5 indicates mild-moderate autism traits, and 37 and higher indicates high autism traits (Schopler et al., 2010).

Table 3. Full model results for the secondary analysis comparing all autistic (coded: 0.5) and NT (coded: -0.5) children tested at the Year 2 time point who passed data cleaning and exclusionary criteria. Nonverbal cognitive ability (grand mean-centered and scaled) was included as a covariate.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.27	0.11	2.45	**.015**
Diagnostic Group	0.03	0.27	0.13	.900
Nonverbal Cognitive Ability	0.09	0.13	0.66	.510

Dispersion Parameter: 7.27

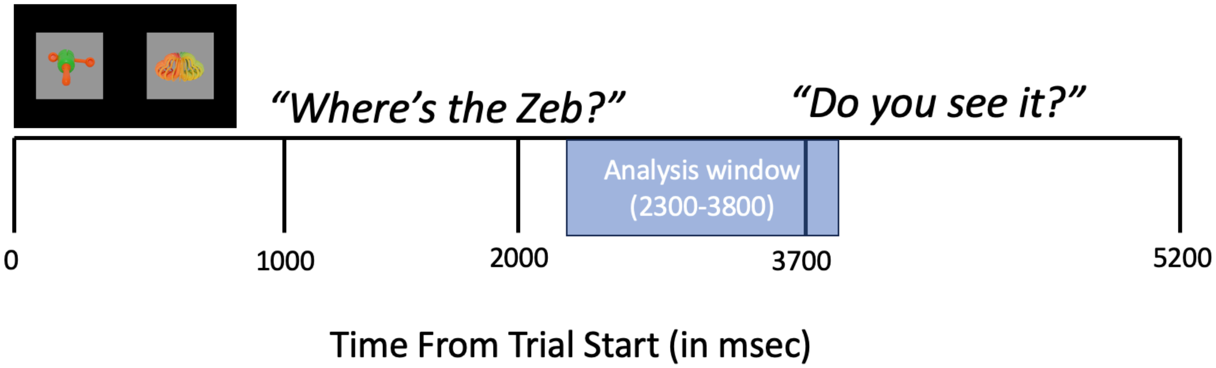


Figure 5. Example test trial schematic. Analysis window shown in blue.

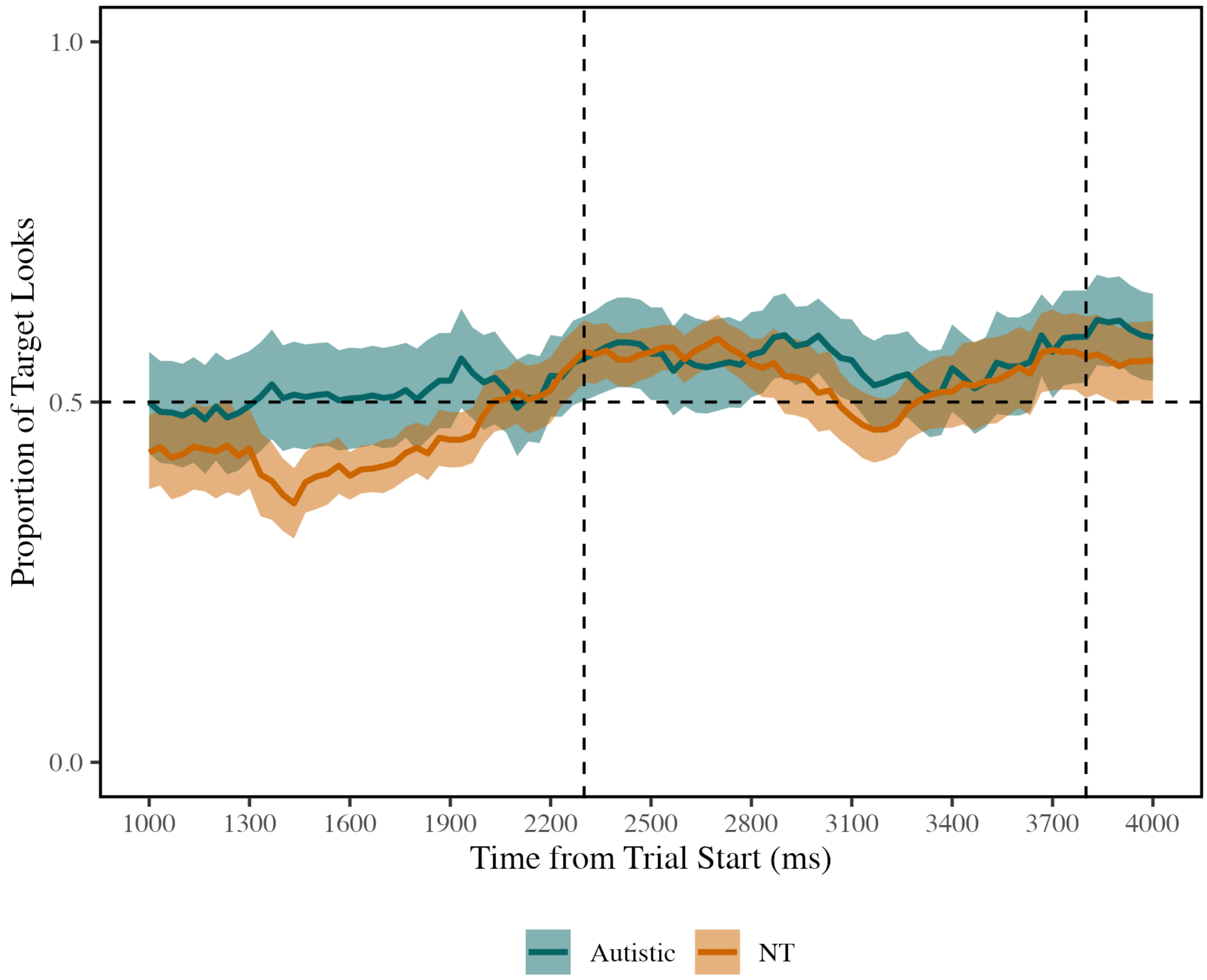


Figure 6. Profile plot showing average proportion looks to target images over time for all autistic (n=19; green) and NT children (n=19; orange) in the nonverbal cognitive ability-matched groups. Line shading indicates one standard error around the mean, averaged across participants. Horizontal dashed line indicates chance (.05). Vertical dashed lines denote the analysis window, 300-1800 ms after target noun onset (which occurred 2000 ms from trial start).

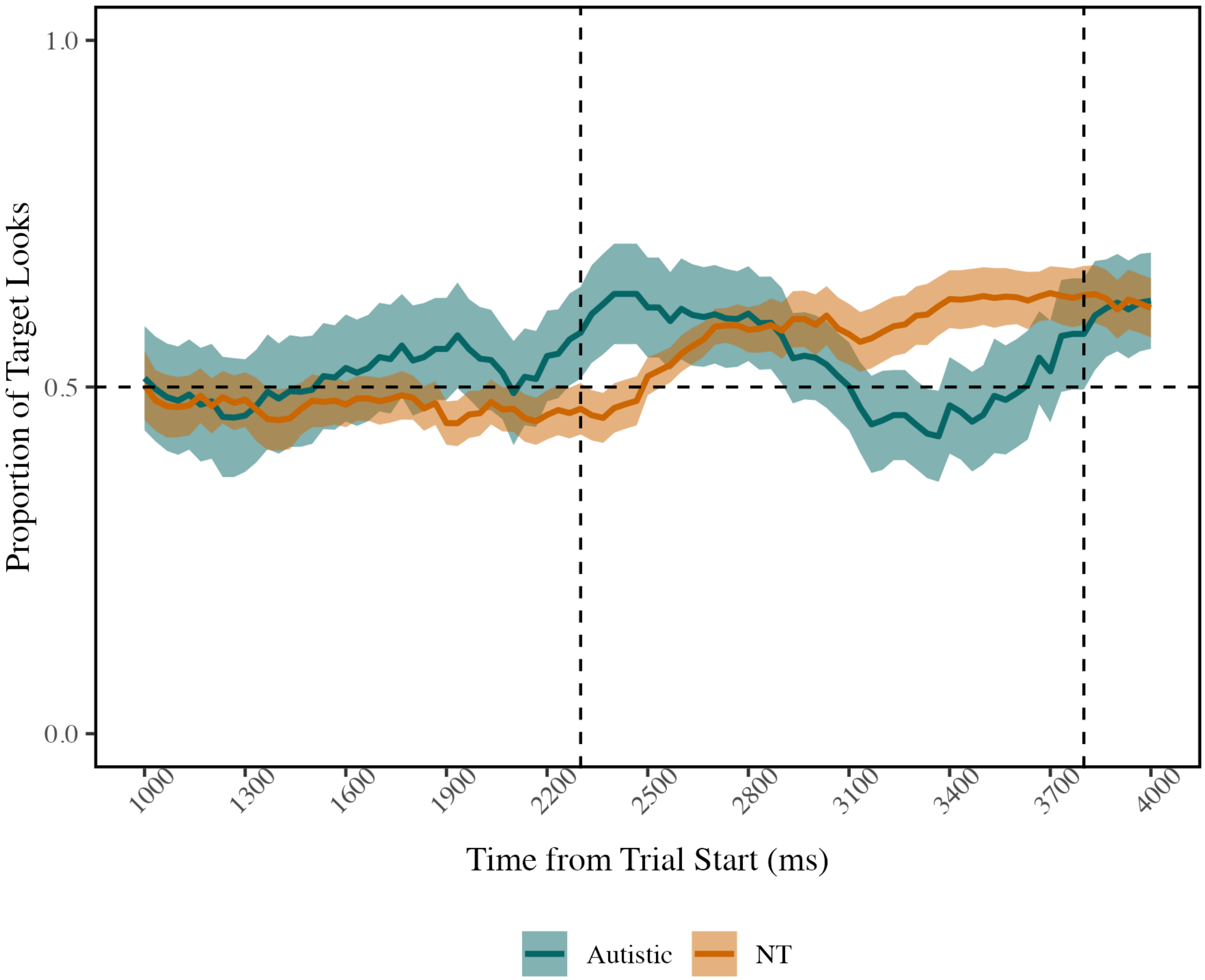


Figure 7. Profile plot showing average proportion looks to target images over time for all autistic (n=17; green) and NT children (n=27; orange) tested at the Year 2 time point who passed data cleaning and exclusionary criteria who were included in the secondary analysis. Line shading indicates one standard error around the mean, averaged across participants. Horizontal dashed line indicates chance (.05). Vertical dashed lines denote the analysis window, 300-1800 ms after target noun onset (which occurred 2000 ms from trial start).

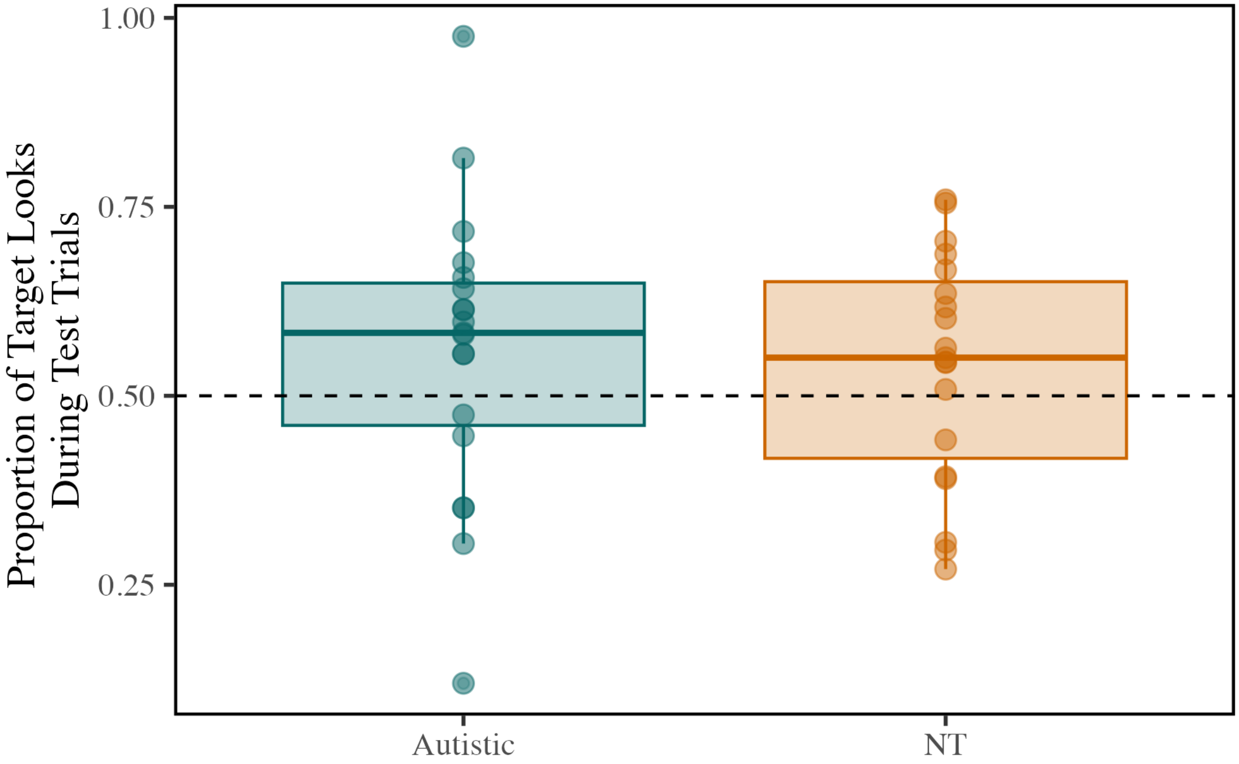


Figure 8. Proportion looks to target images averaged across the analysis window (300-1800 ms after target noun onset) of all included trained (i.e., fast mapping) test trials for nonverbal cognitive ability-matched autistic (n=19; green) and NT children (n=19; orange). Each dot represents one participant's average.

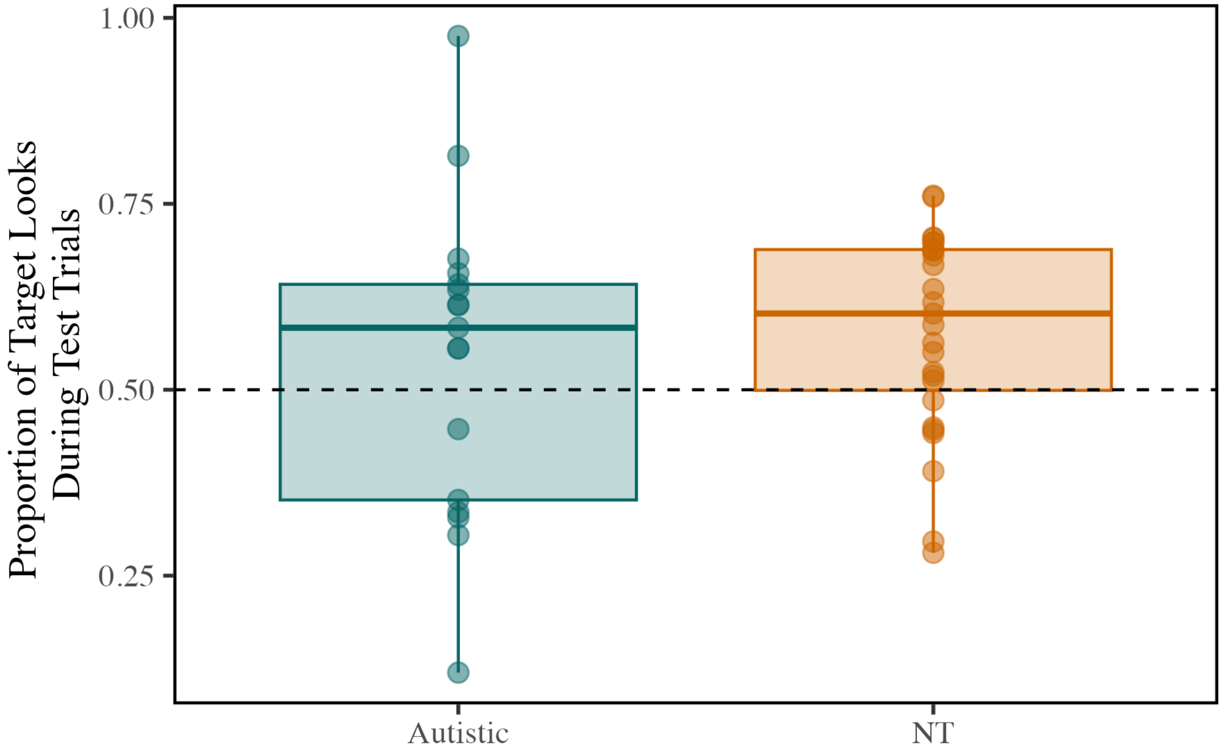


Figure 9. Proportion looks to target images averaged across the analysis window (300-1800 ms after target noun onset) of all included trained (i.e., fast mapping) test trials for all autistic (n=17; green) and NT children (n=27; orange) tested at the Year 2 time point who passed data cleaning and exclusionary criteria who were included in the secondary analysis. Each dot represents one participant's average.

Chapter 4: The Effect of Variable Exemplars in Young Autistic Children's Word Learning

The second study of this dissertation investigated the effect of exemplar variability on word learning acquired through cross-situational word learning (CSL) and ostensive teaching among 3- to 4-year-old autistic children and NT peers. This study employed the data reported from children tested at the Year 2 time point in Chapter 3 (hereafter CSL-NV for cross-situational learning – no variability) in comparison to another experimental eye-gaze task that shared the design of the Chapter 3 task except that each novel word was paired with three exemplar images throughout training that varied in color (hereafter CSL+V for cross-situational learning + variability). The tasks are described in detail in General Methods (Chapter 2). A third eye-gaze task was administered employing an ostensive teaching design with variable exemplars (hereafter O+V for ostensive teaching + variability) in order to determine if the effect of exemplar variability on novel word learning differed depending on whether referents were named explicitly (i.e., O+V) or ambiguously (i.e., CSL+V).

I asked the following research questions:

1. Are 3- to 4-year-old autistic children able to learn novel words through CSL and ostensive teaching with variable exemplars, as measured by accuracy (i.e., proportion of looking to target images greater than chance (.50)?
2. Do 3- to 4-year-old autistic children show differences in novel word learning acquired through CSL with variable exemplars compared to without exemplar variability?
3. Do 3- to 4-year-old autistic children differ from NT peers in their novel word learning acquired through CSL with variable exemplars or in the extent to which learning from CSL differs with and without exemplar variability?

4. Do 3- to 4-year-old autistic children show stronger learning from variable exemplars when they are named explicitly (i.e., ostensive teaching) than when they are named ambiguously (i.e., CSL) and do they differ from NT peers in the extent to which explicit naming improves learning?

Methods

Participants

This study included data from 20 autistic children (5 female; m=44.05 months, SD=3.97, range=36-50) and 30 NT children (15 female, m=32.90 months, SD=2.19, range=30-39). Description of recruitment procedures and exclusionary criteria can be found in General Methods (Chapter 2). All but one NT participant and two autistic participants were administered all three tasks reported in this study (CSL-NV, CSL+V, and O+V). The NT participant missed the CSL-NV task and one of the autistic participants missed the O+V task due to fussiness, and the other autistic child was administered the CSL+V task twice in error instead of O+V. For that child, the first time the task was administered was coded and included in analyses for the CSL+V task only. Most participants (all 20 autistic children and 13 NT children) were tested in these tasks at the Year 2 time point of of a longitudinal parent project, while an additional 17 NT children who were within the Year 2 NT age range were recruited to participate in dissertation tasks only. Four additional NT participants were tested in the dissertation tasks as part of the larger parent study visit protocol at the Year 2 time point but were not included since data analysis for these studies had already begun. Among those tested, not all participants ultimately had data for each study included in analyses due to data cleaning criteria (described below in Data Cleaning). After data cleaning criteria were applied, data were available for 17 autistic children and 27 NT children for the CSL-NV task, 14 autistic children and 29 children for the CSL+V task, and 17 autistic children and 30 NT children for the O+V task. Due to these small sample sizes, it was not feasible to match groups on child characteristic variables. Instead, I included nonverbal cognitive ability (as measured by Developmental Assessment of Young Children – 2nd Edition (DAYC-2) cognitive domain raw scores; Voress & Maddox, 2013) as a covariate in the analyses as I did in Chapter 3. Sample characteristics for participants with data included for the CSL-NV task can be found in Table 2. Sample characteristics for participants with data included for the CSL+V task can be found in Table 4 and participants with data included for the O+V task can be found in Table 5.

Data Analysis

Data Cleaning

Data cleaning procedures were the same as described in Chapter 3. Thirty-three ms video frames in which children were shifting their gaze or looking away from visual stimuli during the analysis window (300-1800 ms after target noun onset in test trials, which occurred 2000 ms from the start of the trial) were removed from analyses. I excluded trials in which children were looking to visual stimuli for less than 50% of the analysis window and excluded participants if they contributed 2 or fewer trained (i.e., fast mapping) trials out of a total of 8 (or 25%). As described in previous chapters, each task included 8 trained (i.e., fast mapping) test trials, 8 untrained (i.e., generalization) test trials, and 2 familiar word test trials (to introduce trial structure and support children's attention). However, an experimental design error rendered generalization test trials uninterpretable. Therefore, analyses included a maximum of 8 trained (i.e., fast mapping) test trials per participant. Across all children with data included for the CSL-NV task (after data cleaning and exclusion), autistic children contributed an average of 5/8 test trials (*SD*=1, range=3-8) and NT children contributed an average of 7/8 test trials (*SD*=1, range=4-8). For the CSL+V task, autistic children contributed 5/8 trials on average (*SD*=2, range=3-8) and included NT children contributed 6/8 trials (*SD*=2, range=3-8). For the O+V task, autistic children contributed an average of 5/8 trials (*SD*=2, range=3-8) and NT children contributed an average of 7/8 trials (*SD*=1, range=4-8).

Analytic Approach

After data cleaning, looks to target (coded 1) and distractor (coded 0) from the analysis window (300-1800 ms after target noun onset) across all trained (i.e., fast mapping) test trials were averaged within each participant for each task. Thus, the outcome variable included one average accuracy measure (i.e., proportion looks to target) per participant for each task in which they contributed sufficient data for inclusion. Raw looking data over time during the analysis window is depicted in the profile plots (Figure 10 and 11) and the aggregated accuracy data used in analyses is pictured in the boxplots (Figure 12 and 13). The first research question investigated whether children were able to learn novel words in the CSL+V task and the O+V task as measured by average accuracy greater than chance (.50). To answer this question, I conducted one-sample, two-tailed t-tests comparing the autistic and NT groups' average accuracy to .50. For the second and third research questions, I examined whether autistic and NT children's novel word learning (as measured by average accuracy) would differ in the CSL+V task compared to the CSL-NV task and whether diagnostic groups differed in the CSL+V task or in the extent of difference in learning between the CSL-NV and CSL+V tasks. To answer these questions, I fit a generalized linear mixed-model using a beta distribution and logit link with the *glmmTMB* package in R (Brooks et al., 2017). This model regressed average accuracy (in terms of log-odds of looking to target due to logit link) on the interaction of diagnostic group (dummy coded: NT=1, autistic=0) and task (dummy-coded: CSL-NV=1, CSL+V=0) with nonverbal cognitive ability (DAYC-2 cognitive domain raw scores, grand mean-centered and scaled) included as a covariate. The model included a random by-subject intercept and by-subject random slope for task, which was the maximal random effects structure (Barr et al., 2013). For the fourth research question, I fit another generalized linear mixed-model using a beta distribution and logit link regressing average accuracy (in terms of log-odds of looking to target due to logit link) on the interaction of diagnostic group (contrasts: NT=1, autistic=0) and task (O+V=1, CSL+V=0) with nonverbal cognitive ability (DAYC-2 cognitive domain raw scores, grand mean-centered and scaled) included as a covariate. This model included a by-subject random intercept only due to convergence problems when a by-subject random slope for task was included. Because the aggregated data used in this model included only one observation per level of the task predictor and task was the only within-subjects predictor in the model, this approach should still provide correct estimates and a 5% type-I error rate according to Brauer and Curtin (2018).

Results

Regarding the first research question, one-sample two-tailed t-tests revealed that autistic children's average proportion looks to target in the CSL+V task significantly exceeded chance, m=.59, $t(13)=2.30$, p=.039 while NT children's did not, m=.47, $t(28)=-0.90$, p=.375, indicating that autistic children as a group learned the novel words in the CSL+V task, while the NT children as a group did not.

The second and third research questions were addressed by the model comparing accuracy on the CSL-NV and CSL+V tasks. The variable of interest for the second research question was task (CSL-NV vs. CSL+V), representing the effect of task for the autistic group based on the dummy coding (i.e., the autistic group was the reference group). This effect was not significant, (β=-0.10, SE=0.24, z=-0.42, p=.672) suggesting that autistic children's looking behavior did not significantly differ between CSL-NV and CSL+V when taking into account all other model terms.

The two variables of interest for the third research question were diagnostic group (representing the effect of group in the CSL+V task based on the dummy coding (i.e., the CSL+V task was the reference group), and the interaction of diagnostic group (NT vs. autistic) and task (CSL-NV vs. CSL+V). The effect of group was significant (β=-0.55, *SE*=0.24, *z*=-2.32, *p*=.021), indicating that the autistic group demonstrated better learning than the NT group in the CSL+V task when taking into account nonverbal cognitive ability and all other model terms. The interaction of diagnostic group and task was not significant (β=0.50, *SE*=0.31, *z*=1.64, *p*=.101) indicating that the extent to which exemplar variability affected novel word learning from CSL did not differ between diagnostic groups when taking all other model terms into account. Full model results can be found in Table 6.

The fourth research question was answered by the model comparing autistic and NT children's looking behavior in the CSL+V and O+V tasks. In this model, the effect of task was not significant, (β=-0.10, *SE*=0.22, *z*=-0.46, *p*=.649) indicating that autistic children's performance did not differ in the CSL+V task compared to the O+V task when taking all other model terms into account. The effect of diagnostic group was significant, (β=-0.58, *SE*=0.27, *z*=-2.16, *p*=.031), indicating again that the autistic group outperformed the NT group in the CSL+V task when covarying nonverbal cognitive ability and taking all other model terms into account. The interaction of task and diagnostic group was significant, (β=0.63, *SE*=0.27, *z*=2.35, *p*=.019). The direction of the effect suggests that, compared to autistic children, NT children showed a larger increase in performance in the O+V task compared to the CSL+V task when taking all other model terms into account. Full model results can be found in Table 7.

Discussion

The purpose of this study was to investigate the impact of exemplar variability in young autistic children's novel word learning acquired through CSL and ostensive teaching and compare that learning to NT peers. The research questions were examined through administration of a series of eye-gaze experiments where children heard novel words while viewing images of their novel referents. In test trials, children were shown the target object alongside a distractor while they heard "*Where's the **gip**?" Do you like/see it?*" Children's eye movements were recorded and then coded offline frame-by-frame for the outcome measure. The CSL-NV task design followed a classic CSL paradigm, described in more detail in previous chapters. The CSL+V task added exemplar variability to the CSL paradigm, pairing each novel word with three exemplar images presented during training that varied in color. The O+V task also paired three exemplar images with each novel word during training, but presented each novel word-object pairing individually in each trial such that each object was named explicitly. This task therefore allowed us to examine whether explicit naming improved novel word learning from exemplar variability over ambiguous naming (i.e., CSL).

Because previous research has suggested that autistic children do not show a shape bias in early development (Hartley & Allen, 2014; Potrzeba et al., 2015; Tek et al., 2008), possibly due to underlying differences such as difficulty abstracting category-relevant properties across exemplars (Tek & Naigles, 2017) which may be related to fragmented categorical representations competing for processing (Tovar et al., 2020), I expected autistic children to have more difficulty learning words when training exemplars varied perceptually (i.e., the CSL+V task) compared to when a novel word was paired with only a single exemplar (i.e., the CSL-NV task). I further expected autistic children to have more difficulty learning words from variable exemplars (i.e., in the CSL+V task) than NT peers. Finally, I expected that children might be more successful learning novel words from variable exemplars when they were named explicitly (i.e., in the O+V task) than ambiguously (i.e., in the CSL+V task). Study results were surprising; contrary to expectations, autistic children as a group demonstrated above-chance accuracy in the CSL+V task, indicating that they were on average able to learn novel words from variable exemplars. By contrast, the NT children as a group did not significantly differ from chance in the CSL+V task. This finding was likewise surprising as evidence in the literature indicates that variability is not detrimental to NT children's word learning at least in terms of immediate referent selection (i.e., fast mapping) for words referring to concrete objects, and is even beneficial in terms of retention and generalization of those words (Crespo & Kaushanskaya, 2021; Perry et al., 2010; Twomey et al., 2014, 2018). Further, when diagnostic groups were directly compared, autistic children showed significantly better learning in the CSL+V task compared to NT peers when accounting for differences in nonverbal cognitive ability as a covariate and all other model terms (Table 6). While model results did not yield a significant difference between CSL-NV and CSL+V task performance for the autistic group in this study, it is notable that the same children did not show above-chance accuracy in the CSL-NV task as reported in the secondary analysis of Chapter 3. The lack of differences found between tasks when directly compared might be attributed to the wide variability in autistic children's performance in the CSL-NV task such that there was considerable overlap between tasks (Figure 12) or to the relatively small sample size. This pattern of findings contrasts with that of the NT children, who at the Year 2 time point showed above-chance accuracy in the CSL-NV task (as reported in the secondary analysis of Chapter 3) but not in the CSL+V task.

Also contrary to expectations, explicitly labeling novel words referring to variable exemplars (i.e., the O+V task) did not improve learning over the CSL+V task for autistic children. In fact, autistic children did not differ from chance in O+V task accuracy. The NT children were above chance in the O+V task, and in comparison to the autistic group showed a larger increase in learning from exemplar variability in the O+V task compared to the CSL+V task. These findings align more closely with the majority of the extant literature suggesting that exemplar variability at a minimum is not deleterious to NT children's learning of words referring to concrete objects (Crespo & Kaushanskaya, 2021; Perry et al., 2010; Twomey et al., 2014, 2018). However, it does appear that the combination of exemplar variability *and* referential ambiguity was particularly difficult for NT children in this sample, who were otherwise able to learn novel words with above-chance accuracy in both a CSL task without variability (CSL-NV) and an explicit labeling task with variability (O+V). This finding might be expected based on the assumption that combining a CSL task with variable exemplars would increase the difficulty in comparison to either a basic CSL task or exemplar variability alone. For example, 10-month-old NT children who were able to learn word-object mappings from a single exemplar failed to show evidence of learning when presented with variable exemplars during training, which the researchers attributed to the cognitive load of the multiple exemplar condition in very young children (Taxitari et al., 2020). As another example, learning verbs is more difficult than learning nouns for young children (e.g., Gentner, 1982, 2006; Gentner & Boroditsky, 2001) and variability in actors during novel verb training was found to be detrimental in comparison to one actor for 2.5- and 3-year-old NT children (Maguire et al., 2008).

Interestingly, the opposite pattern held for the autistic children. They demonstrated above-chance learning *only* in the task including both exemplar variability and a CSL paradigm

(CSL+V). While surprising, there are a few possible explanations for this pattern of findings. First, the design of this study may have allowed children to rely on different mechanisms to learn words than those involved in shape bias studies. For example, in the fast-mapping test trials, children only had to associate a novel word with one of the exemplars they had already seen, not extend the learned novel label to a new object based on either shape or some other characteristic (e.g., function or color), possibly allowing children to rely more heavily on simple associative learning. Or, perhaps how narrowly exemplars varied during training (i.e., in color only) was not sufficient to elicit the expected effect in autistic children. According to computational modeling by Tovar et al. (2020), findings of a lack of shape bias in autistic children may be attributed to more fragmented, context-bound category representations competing for processing of incoming stimuli. If the within-lexical category exemplars in the present study were too similar perceptually, they may not have produced the kind of interference described by Tovar et al. (2020), and instead been perceived as essentially the same image, again allowing children to rely on associative learning. However, these explanations do not account for the *superior* performance by autistic children in the CSL+V task. While the within-lexical category exemplars varied relatively narrowly (i.e., on only one perceptual dimension), that the colors of the images on screen changed more frequently throughout the training trials may have added to the task's overall perceptual salience and therefore enhanced autistic children's visual attention. That is, autistic children may have preferentially attended to this higher salience task to a larger extent than NT peers. Previous findings offer some support for this view; for example, autistic individuals' fixations to visual scenes have revealed a greater pixel-level saliency bias (e.g., color, intensity, and orientation) but reduced bias at the object-level (e.g., size, complexity, solidity) and semantic-level (e.g., face, emotion, motion, operability) relative to NT comparisons

(Wang et al., 2015). That the objects changed color frequently throughout the task may have therefore added to the saliency of the CSL+V task specifically for autistic children. This explanation would also align with a previous eye-gaze study from our research group which found that more perceptually salient distractor images of familiar objects drew autistic (but not NT) children's attention away from named but less perceptually salient images of familiar objects (Venker et al., 2021).

However, another study found that autistic children's novel referent selection was actually hindered by high perceptual salience compared to low perceptual salience (Venker et al., 2022). The authors attributed this effect to more difficulty visually disengaging from high salience distractor images to look to named high salience target images in referent selection test trials compared to low salience trials, often referred to as "sticky" attention. It is worth noting that children saw only one image at a time during training in that study, more similar to the O+V task where autistic children also did not show above-chance accuracy. Perhaps moving from one image on screen at a time during training to two images during test encouraged more visual exploration between high salience targets and distractors among autistic children. By contrast, the CSL+V task included two images on screen at a time through both training and test, possibly allowing children more time to habituate to the complexity of the visual array. The training phase of the CSL+V task also moved relatively quickly (each trial lasted ~3.5 seconds), there was a half-second pause between trials, and images appeared on both sides of the screen across trials. Therefore, even if autistic children did have "sticky" attention for high salience images, the task provided ample opportunities to shift their gaze and view different images throughout training. Additionally, children only received three exposures to the novel word while viewing each novel object during training in Venker et al. (2022). In the present study, children had 10 exposures to each novel word-object pairing in each task. It is possible that sufficient exposure allows autistic children to overcome lower-level perceptually-driven looking preference for higher salience and even result in better learning, especially if interest or attention were better sustained as a result. Such a possibility could have important implications for language intervention and should be explored further in future work.

If autistic children's superior performance in the CSL+V task was due to increased visual attention during training, one metric that may provide evidence is the amount of time children were shifting or looking away from (i.e., missed) the visual stimuli during training trials. I calculated the average percent of frames children missed within a window during training trials between the onset of the first spoken novel word and the offset of the second spoken novel word (500-3700 ms after training trial start). Across all autistic children who had available data (but before exclusion based on data cleaning criteria; *n*=20 for each task), children missed 34.53% of the training trial window on average (*SD*=17.52, range=8.66-71.44) for the CSL-NV task. For the CSL+V task, children missed 29.88% on average (SD=13.09, range=8.45-55.05). While less in absolute terms, a paired *t*-test revealed this was not a statistically significant difference, *t*(19)=1.34, *p*=.194. Therefore, it does not appear that autistic children were more attentive to training trials in the CSL+V task relative to the CSL-NV task, at least in terms of frames in which they looked to visual stimuli. However, it should be noted that this is a relatively course measure of attention as the response codes only indicated whether children were or were not directing their gaze to one of the two areas of interest (AOIs; i.e., the grey boxes around each image). This measure did not account for visual exploration of image features within the AOI, time spent looking at each AOI, or number of shifts between the two AOIs. It is possible that simply looking in the direction of named referents is not equivalent to fully engaging in a word learning opportunity. As further evidence that there may be a disconnect between attention during training as measured by time spent looking to visual stimuli and resulting learning, the NT group on average missed significantly less of the training trials in the CSL+V task compared to the autistic group based on a two-sample two-tailed *t*-test (*t*(29.88)=-3.64, *p*=.001) despite an average accuracy not differing from chance. The distinction between visual attention based on stimulus salience and attention based on learned associations has been explored in a previous CSL study with young NT children (Smith & Yu, 2013). By manipulating trial order such that visual attention based on local stimulus novelty could be distinguished from attention based on statistical regularities, they found that children who learned the words in the task showed habituation to the local novelty effects relative to the non-learners, who showed sensitivity to stimulus novelty throughout the task. Unlike the task in Smith and Yu (2013), the novelty and salience of the exemplars presented during training in the CSL+V task should have been relatively consistent across trials. The same within-lexical category exemplar never repeated from one naming instance to the next, and there were the same number of exemplars for each novel word (see Appendix A and B). Therefore, stimulus novelty/salience should be less likely to compete with the unfolding statistical regularities for attention in the CSL+V task. If salience remained at a moderate, consistent level across trials, perhaps it acted instead as an attentional aid to the regularities informing the correct word-referent mappings specifically for children who are more sensitive to low-level perceptual salience. More fine-grained analysis of autistic children's attention to and interest in high and low salience word learning tasks and resulting learning may be a fruitful area for future research.

One final remaining question regarding the findings of the present study is why autistic children were above chance in the CSL+V task but not the O+V task, which should have been easier as each object was labeled explicitly one at a time. One possibility was discussed above; perhaps viewing only one object per training trial in the O+V task encouraged “sticky” attention during test because children were not already habituated to seeing two highly salient objects side-by-side as they were in the CSL+V task. If autistic children do prefer low-level perceptual salience, they may have been more driven to visually explore the images than to fixate based on spoken labels. Another possibility is the order of task administration. Children were randomly assigned to one of two administration protocols. In both protocols, the CSL+V task was administered after at least two others but never last, while the O+V task was administered after only one other task in the first protocol and last in the second protocol. It is likely, therefore, that at least some children were affected by fatigue during administration of the O+V task to a greater extent than the CSL+V task. It is worth noting that most autistic children were individually more accurate than chance in both tasks (Figure 13). Indeed, the autistic children’s median average accuracy in both tasks (represented by the bold line green lines in Figure 13) are very similar (*Mdn*=.60 in CSL+V and *Mdn*=.59 in O+V). However, the child-level averages in the O+V task appear to skew lower than the CSL+V task which likely brought down the average accuracy at the group level.

Limitations and Future Directions

The present results should be considered in the context of several potential limitations. First, applying data cleaning criteria and the fact that some children were not able to complete all experimental tasks resulted in unequal numbers of children with data available for analysis for each task. I chose to analyze all the available data and compared across tasks using mixed-effects models under the assumption that the data are missing at random (MAR; e.g., Pugh et al., 2021). I am assuming MAR for the purposes of this analysis because the data are largely missing due to the study design choice to exclude at the task level based on number of trials contributed, one participant had missing O+V data due to experimenter error, and only one participant each for the CSL-NV and O+V were too fussy to continue the protocol. I chose to exclude children at the task level if they contributed two or fewer trials after data cleaning under the assumption that one to two trials would be too noisy to produce a reliable average accuracy measure. Therefore, while this decision reduced some bias in the outcome measure, I cannot entirely rule out the possibility that children who contributed less data and were therefore excluded for a given task differed in some way that was not accounted for by the available data and included covariates. This limitation may be particularly relevant for the CSL+V task, as only 14 autistic children had sufficient data for inclusion based on the criterion used (compared to 17 for the other two tasks), and this was the only task in which autistic children as a group were significantly above chance. For example, among many possible sensory profiles, some autistic children may have more sensory-seeking tendencies while others may have more sensory-avoidant tendencies (e.g., Little et al., 2017). Perhaps sensory-avoidant children found the perceptual experience of the CSL+V task too overwhelming and looked away more often. Excluded children's inattention to the task could also be related to abilities in other domains. For example, Venker (2019) found that the autistic children excluded from analyses due to insufficient contributed data had the lowest language skills. Additionally, as in all studies in this dissertation, I acknowledge that the present findings may not generalize to the broader population given that the sample was primarily composed of white, non-Hispanic or Latino children from higher socioeconomic status backgrounds (see Table 4 and 5 for participant characteristics).

Another limitation of the present study that should be noted was the use of a covariate approach to account for diagnostic group differences in nonverbal cognition. As discussed in

Chapter 3, I chose to covary nonverbal cognitive ability for two reasons: first, because domain-general mechanisms at least partially support statistical learning (Frost et al., 2015; Kirkham et al., 2002; Krogh et al., 2013) and CSL in particular is related to multiple cognitive domains in NT children (Vlach & DeBrock, 2017). Second, the larger parent study under which most participants were originally recruited was designed to match groups on nonverbal cognition. As such, the range of scores on the nonverbal cognitive measure between groups were more similar than other measures. While commonly used, the validity of the covariate approach in accounting for group differences has been questioned in cases where the covariate is meaningfully related to group (Dennis et al., 2009; Miller & Chapman, 2001). While a large subset of autistic individuals have cognitive abilities within or above the average range for their age, current estimates suggest 37.9% of autistic 8-year-olds have an intellectual disability (at least among those with data available; Maenner et al., 2023). Thus, it is arguable whether cognitive ability is meaningfully related to autism diagnosis and significant diagnostic group effects found in this study should be cautiously interpreted. One might also consider the role of other child characteristics that differ between these groups even when accounting for nonverbal cognitive ability in analyses, such as receptive language and chronological age (Table 4 and Table 5). The NT children were recruited at an intentionally younger age than the autistic children in order to match on nonverbal cognition, and therefore diagnostic group is highly confounded with chronological age in this sample. Thus, if I were to include age in the models, the variance accounted for by age would likely absorb much of the group effect. However, results of the present study should be considered within the context that the autistic children were older and might have some developmental advantage over the NT group. Future studies including an age- and cognitive ability-matched NT sample might clarify this question. Additionally, given that these tasks are linguistic in nature, it is reasonable to question the role of receptive language abilities. However, the fact that the autistic children outperformed the NT children in the CSL+V task despite significantly lower receptive language abilities as a group suggests that receptive language would not account for the diagnostic group effect. Individual child characteristics as predictors of performance within the autistic group, including age and receptive language, will be examined in Chapter 5.

Finally, it is important to acknowledge that learning a word involves more than immediate mapping of labels to referents. Unfortunately, the experimental design error that prevented analysis of generalization trials leaves open the important question of whether exemplar variability might differentially affect autistic children's willingness to extend a label to a novel exemplar, and the present study did not include retention test trials to understand how children's learning from CSL and exemplar variability evolves over time. Just because autistic children appeared to demonstrate encoding of novel words would not necessarily preclude forgetting after a delay or if learning opportunities were spaced apart in time. The importance of examining the impact of exemplar variability on generalization and retention of word learning in autistic children will be elaborated in General Discussion (Chapter 6) and should be examined in future studies.

Conclusions

This study aimed to investigate the impact of exemplar variability on young autistic and NT children's word learning in a CSL paradigm and compare that learning to a CSL task without variability and an ostensive teaching task with variability to better understand how young autistic children learn in ambiguous (i.e., CSL) and perceptually variable contexts. Contrary to expectations, results suggest that autistic children were able to learn novel words from a CSL task with exemplar variability and performed better than NT peers when nonverbal cognition was covaried. Compared to the autistic group, the NT group showed a larger increase in performance from the CSL task with variability to the ostensive teaching task with variability, suggesting that NT children benefitted from explicit 1:1 naming in the context of exemplar variability to a greater degree than the autistic children. Given the inclusion of nonverbal cognitive ability as a covariate, significant group effects should be interpreted cautiously. While surprising, these findings may highlight the role of perceptual salience in autistic children's word learning and should be investigated further including generalization and retention test items.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Brooks, M., E., Kristensen, K., Benthem, K., J. ,van, Magnusson, A., Berg, C., W., Nielsen, A., Skaug, H., J., Mächler, M., & Bolker, B., M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378. https://doi.org/10.32614/RJ-2017-066

Crespo, K., & Kaushanskaya, M. (2021). Is 10 Better than 1? The Effect of Speaker Variability on Children’s Cross-Situational Word Learning. *Language Learning and Development*, *17*(4), 397–410. https://doi.org/10.1080/15475441.2021.1906680

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*(3), 331–343. https://doi.org/10.1017/S1355617709090481

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Gentner, D. (1982). *Why Nouns are Learned Before Verbs: Linguistic Relativity Versus Natural Partitioning. Technical Report No. 257.*

Gentner, D. (2006). Why Verbs Are Hard to Learn. In K. A. Hirsh-Pasek & R. M. Golinkoff, *Action Meets Word* (1st ed., pp. 544–564). Oxford University PressNew York. https://doi.org/10.1093/acprof:oso/9780195170009.003.0022

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (1st ed., pp. 215–256). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.010

Hartley, C., & Allen, M. L. (2014). Brief report: Generalisation of word-picture relations in children with autism and typically developing children. *Journal of Autism and Developmental Disorders*, *44*(8), 2064–2071. https://doi.org/10.1007/s10803-014-2074-1

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Krogh, L., Vlach, H., & Johnson, S. (2013). Statistical Learning Across Development: Flexible Yet Constrained. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00598

Little, L. M., Dean, E., Tomchek, S. D., & Dunn, W. (2017). Classifying sensory profiles of children in the general population. *Child: Care, Health and Development*, *43*(1), 81–88. https://doi.org/10.1111/cch.12391

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … Shaw, K. A. (2023). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *MMWR Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*(4), 628–634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48. https://doi.org/10.1037/0021-843X.110.1.40

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189

Potrzeba, E., Fein, D., & Naigles, L. (2015). Investigating the shape bias in typically developing children and children with autism spectrum disorders. *Frontiers in Psychology*, *6*(MAR), 1–12. https://doi.org/10.3389/fpsyg.2015.00446

Pugh, S. L., Brown, P. D., & Enserro, D. (2021). Missing repeated measures data in clinical trials. *Neuro-Oncology Practice*, *9*(1), 35–42. https://doi.org/10.1093/nop/npab043

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale—Second Edition (CARS-2)*. Western Psychological Services.

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3/4), 591–611.

Smith, L. B., & Yu, C. (2013). Visual Attention Is Not Enough: Individual Differences in Statistical Word-Referent Learning in Infants. *Language Learning and Development*, *9*(1), 25–49. https://doi.org/10.1080/15475441.2012.707104

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3)*. Pearson.

Taxitari, L., Twomey, K. E., Westermann, G., & Mani, N. (2020). The Limits of Infants' Early Word Learning. *Language Learning and Development*, *16*(1), 1–21. https://doi.org/10.1080/15475441.2019.1670184

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, *1*(4), 208–222. https://doi.org/10.1002/aur.38

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

Tovar, Á. E., Rodríguez-Granados, A., & Arias-Trejo, N. (2020). Atypical shape bias and categorization in autism: Evidence from children and computational simulations. *Developmental Science*, *23*(2), 1–14. https://doi.org/10.1111/desc.12885

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the Right Noises: Background Variability Helps Early Word Learning. *Cognitive Science*, *42*, 413–438. https://doi.org/10.1111/cogs.12539

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's More Like It: Multiple Exemplars Facilitate Word Learning. *Infant and Child Development*, *23*, 105–122. https://doi.org/10.1002/icd

Venker, C. E., Mathée, J., Neumann, D., Edwards, J., Saffran, J., & Ellis Weismer, S. (2021). Competing Perceptual Salience in a Visual Word Recognition Task Differentially Affects Children With and Without Autism Spectrum Disorder. *Autism Research*, *14*(6), 1147–1162. https://doi.org/10.1002/aur.2457

Venker, C. E., Neumann, D., & Aladé, F. (2022). Visual perceptual salience and novel referent selection in children with and without autism spectrum disorder. *Autism & Developmental Language Impairments*, *7*, 239694152210854. https://doi.org/10.1177/23969415221085476

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230. https://doi.org/10.1016/j.jml.2016.10.001

Voress, J. K., & Maddox, T. (2013). *Developmental Assessment of Young Children-Second Edition (DAYC-2)*. PRO-ED.

Wang, S., Jiang, M., Duchesne, X. M., Laugeson, E. A., Kennedy, D. P., Adolphs, R., & Zhao, Q. (2015). Atypical Visual Saliency in Autism Spectrum Disorder Quantified through Model-Based Eye Tracking. *Neuron*, *88*(3), 604–616. https://doi.org/10.1016/j.neuron.2015.09.042

Wilcoxon, F. (1945). Some Uses of Statistics in Plant Pathology. *Biometrics Bulletin*, *1*(4), 41–45. https://doi.org/10.2307/3002011

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Tables and Figures

Table 4. Participant characteristics and demographic information for children with data included for the CSL+V task.

	Autistic Group (n=14)	NT Group (n=29)	Group Comparison
	Mean (SD) Median Range	Mean (SD) Median Range	
Participant Characteristics			
Age (Months)	44.07 (3.79) 45 36-50	32.97 (2.20) 33 30-39	*p*<.001
Receptive Language			
Raw Score	25.50 (11.97) 19 11-53	42.15 (8.20) 41.5 30-65	*p*<.001
Standard Score	64.57 (22.57) 50 50-125	115.92 (14.36) 117.5 87-142	*p*<.001
Expressive Language			
Raw Score	27.14 (8.56) 24 13-44	38.42 (4.24) 37.5 33-49	*p*<.001
Standard Score	69.79 (15.15) 62.5 50-103	111.00 (9.58) 108 97-129	*p*<.001
Cognitive Ability			
Raw Score	38.14 (8.93) 35 28-57	44.90 (4.62) 44 36-60	*p*<.001
Standard Score	76.71 (12.24) 73.5 56-104	106.10 (7.59) 105 91-122	*p*<.001

Adaptive Behavior	67.93 (7.12) 67 55-78	102.24 (12.03) 102 80-131	*p*<.001
Level of Autism Traits	35.50 (4.78) 35 26.5-47	—	—
Demographic Information			
Maternal Education (Years)	14.43 (2.71) 14 11-19	17.34 (2.73) 16 12-26	*p*=.005
Race	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 1 Black or African American 3 More than one race 10 White	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 1 More than one race 28 White	
Ethnicity	1 Hispanic or Latino 13 Not Hispanic or Latino 0 Unknown/Not Reported	0 Hispanic or Latino 28 Not Hispanic or Latino 1 Unknown/Not Reported	

Note. Group differences were quantified by two-sample, two-tailed t-tests if both groups were normally distributed as measure by Shapiro-Wilk normality test (Arunachalam & Luyster, 2018) or Wilcoxon rank sum tests (Wilcoxon, 1945) with continuity correction if either was not. Resulting p-values are reported in the third column. Receptive language was measured by PLS-5 Auditory Comprehension subtest scores (Zimmerman et al., 2011). Expressive language was measured by PLS-5 Expressive Communication subtest scores. Three NT participants did not complete the PLS-5. Cognitive ability was measured by DAYC-2 Cognitive Domain scores. Adaptive Behavior was measured by Vineland-III Adaptive Behavior Composite Scores (Sparrow et al., 2016). Level of autism traits was measured by CARS-2 total scores, where 15-29.5 indicates minimal-no autism traits, 30-36.5 indicates mild-moderate autism traits, and 37 and higher indicates high autism traits (Schopler et al., 2010).

Table 5. Participant characteristics and demographic information for children with data included for the O+V task.

	Autistic Group (n=17)	NT Group (n=30)	Group Comparison
	Mean (SD) Median Range	Mean (SD) Median Range	
Participant Characteristics			
Age (Months)	44.71 (3.80) 46 36-50	32.90 (2.19) 32.5 30-39	p<.001
Receptive Language			
Raw Score	25.71 (11.54) 26 11-53	42.07 (8.05) 41 30-65	p<.001
Standard Score	65.00 (21.48) 57 50-125	115.89 (14.08) 117 87-142	p<.001
Expressive Language			
Raw Score	27.94 (7.69) 24 18-44	38.41 (4.16) 38 33-49	p<.001
Standard Score	70.88 (15.14) 63 52-103	111.00 (9.40) 108 97-129	p<.001
Cognitive Ability			
Raw Score	38.24 (8.27) 36 28-57	44.97 (4.56) 44 36-60	p<.001
Standard Score	75.94 (12.48) 74 56-104	106.23 (7.50) 105 91-122	p<.001
Adaptive Behavior	68.82 (7.29) 69 53-79	102.67 (12.05) 102.5 80-131	p<.001

Level of Autism Traits	35.21 (3.87) 35 26.5-43.5	—	—
Demographic Information			
Maternal Education (Years)	13.82 (2.35) 12 11-18	17.30 (2.69) 16 12-26	*p*<.001
Race	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 5 More than one race 12 White	0 American Indian/Alaska Native 0 Asian 0 Native Hawaiian/Pacific Islander 0 Black or African American 1 More than one race 29 White	
Ethnicity	3 Hispanic or Latino 14 Not Hispanic or Latino 0 Unknown/Not Reported	0 Hispanic or Latino 29 Not Hispanic or Latino 1 Unknown/Not Reported	

Note. Group differences were quantified by two-sample, two-tailed t-tests if both groups were normally distributed as measure by Shapiro-Wilk normality test (Shapiro & Wilk, 1965) or Wilcoxon rank sum tests (Wilcoxon, 1945) with continuity correction if either was not. Resulting p-values are reported in the third column. Receptive language was measured by PLS-5 Auditory Comprehension subtest scores (Zimmerman et al., 2011). Expressive language was measured by PLS-5 Expressive Communication subtest scores. Three NT participants did not complete the PLS-5. Cognitive ability was measured by DAYC-2 Cognitive Domain scores. Adaptive Behavior was measured by Vineland-III Adaptive Behavior Composite Scores (Sparrow et al., 2016). Level of autism traits was measured by CARS-2 total scores, where 15-29.5 indicates minimal-no autism traits, 30-36.5 indicates mild-moderate autism traits, and 37 and higher indicates high autism traits (Schopler et al., 2010).

Table 6. Full model results for the analysis comparing autistic (coded: 0) and NT (coded: 1) children's performance in the CSL+V (coded: 0) and the CSL-NV (coded: 1) tasks. Nonverbal cognitive ability (grand mean-centered and scaled) was included as a covariate.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.41	0.19	2.16	**.031**
Diagnostic Group	-0.55	0.24	-2.32	**.021**
Task	-0.10	0.24	-0.42	.671
Nonverbal Cognitive Ability	0.12	0.09	1.31	.191
Diagnostic Group x Task	0.50	0.30	1.64	.101

Dispersion Parameter: 9.51

Table 7. Full model results for the analysis comparing autistic (coded: 0) and NT (coded: 1) children's performance in the CSL+V (coded: 0) and the O+V (coded: 1) tasks. Nonverbal cognitive ability (grand mean-centered and scaled) was included as a covariate.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.40	0.21	1.87	.061
Diagnostic Group	-0.58	0.27	-2.16	**.031**
Task	-0.10	0.22	-0.46	.649
Nonverbal Cognitive Ability	0.09	0.12	0.84	.400
Diagnostic Group x Task	0.63	0.27	2.35	**.019**

Dispersion Parameter: 11.8

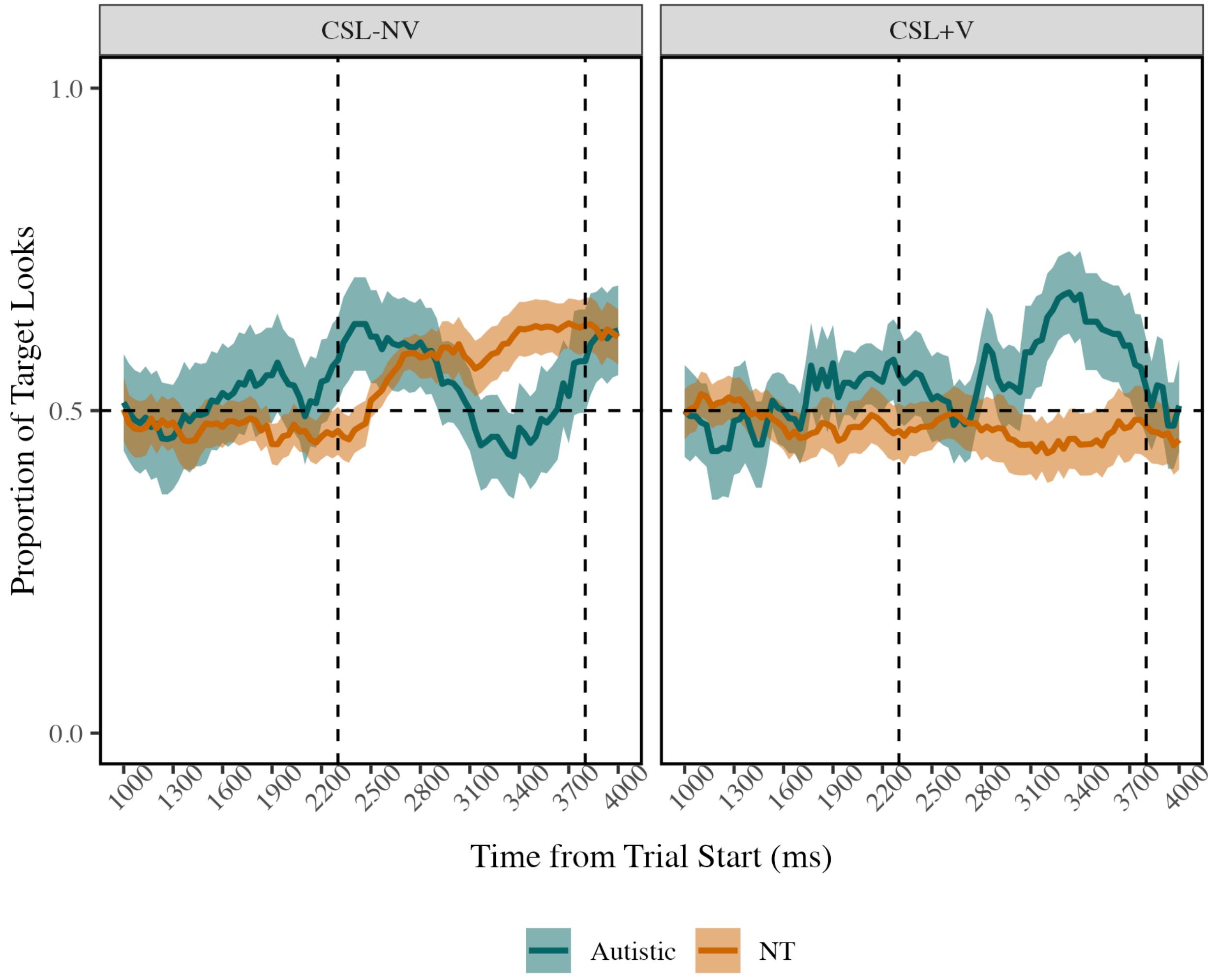


Figure 10. Profile plot showing average proportion looks to target images over time for all autistic (green) and NT (orange) children included in analyses for the CSL-NV task (left) and CSL+V task (right). Line shading indicates one standard error around the mean, averaged across participants. Horizontal dashed line indicates chance (.05). Vertical dashed lines denote the analysis window, 300-1800 ms after target noun onset (which occurred 2000 ms from trial start).

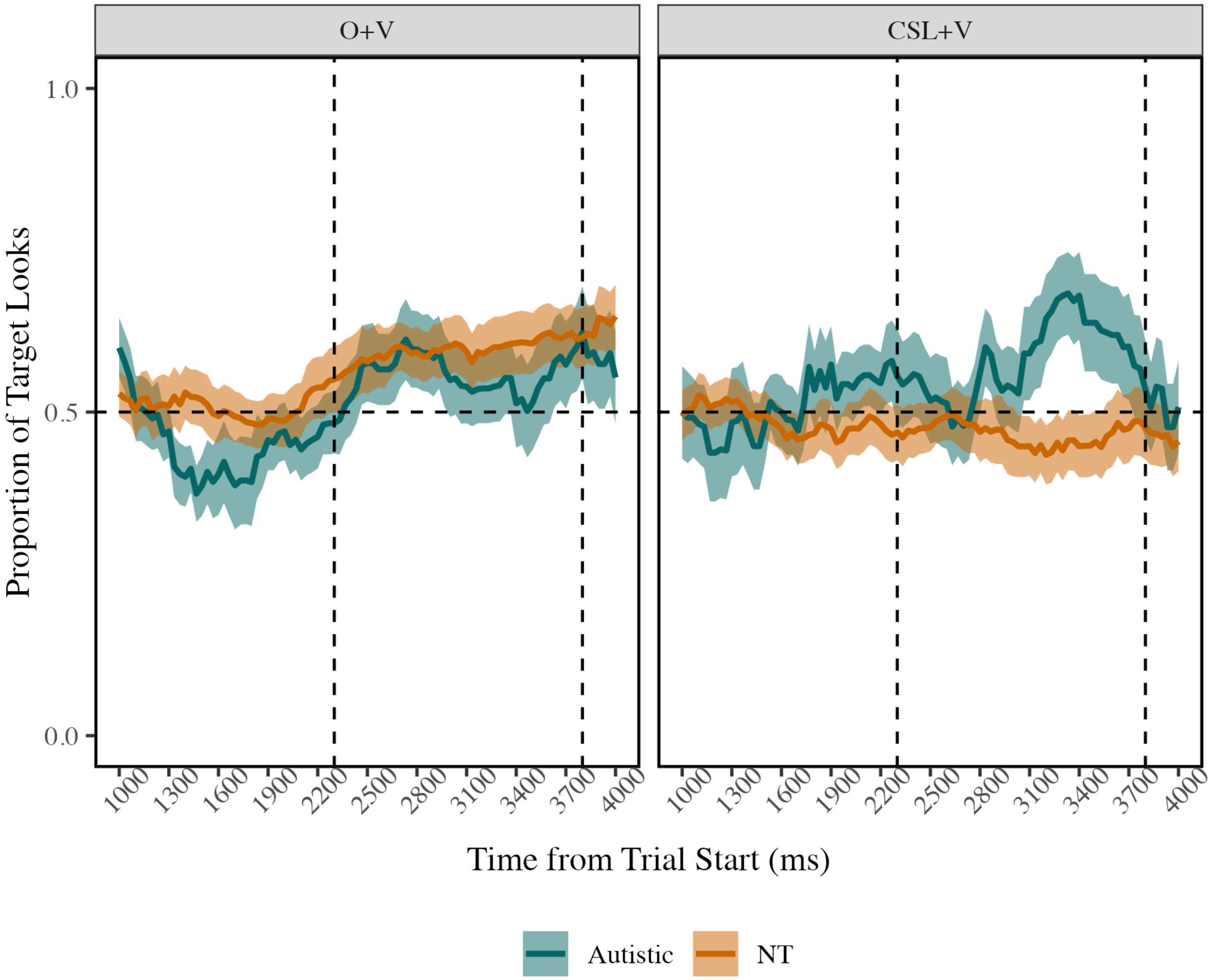


Figure 11. Profile plot showing average proportion looks to target images over time for all autistic (green) and NT (orange) children included in analyses for the O+V task (left) and CSL+V task (right). Line shading indicates one standard error around the mean, averaged across participants. Horizontal dashed line indicates chance (.05). Vertical dashed lines denote the analysis window, 300-1800 ms after target noun onset (which occurred 2000 ms from trial start).

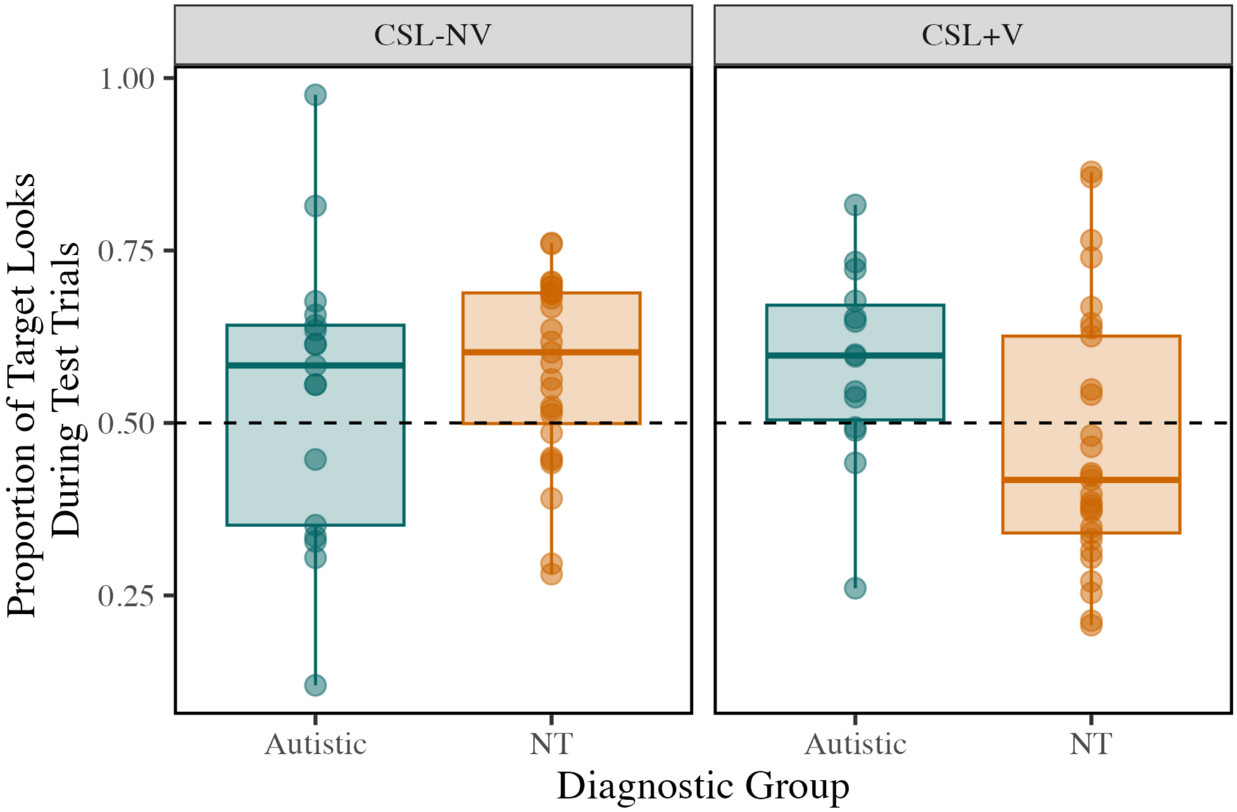


Figure 12. Proportion looks to target images averaged across the analysis window (300-1800 ms after target noun onset) during trained (i.e., fast mapping) test trials for NT (green) and autistic children (orange) with data included in analyses for the CSL-NV task (left) and the CSL+V task (right). Each dot represents one participant's average in that task.

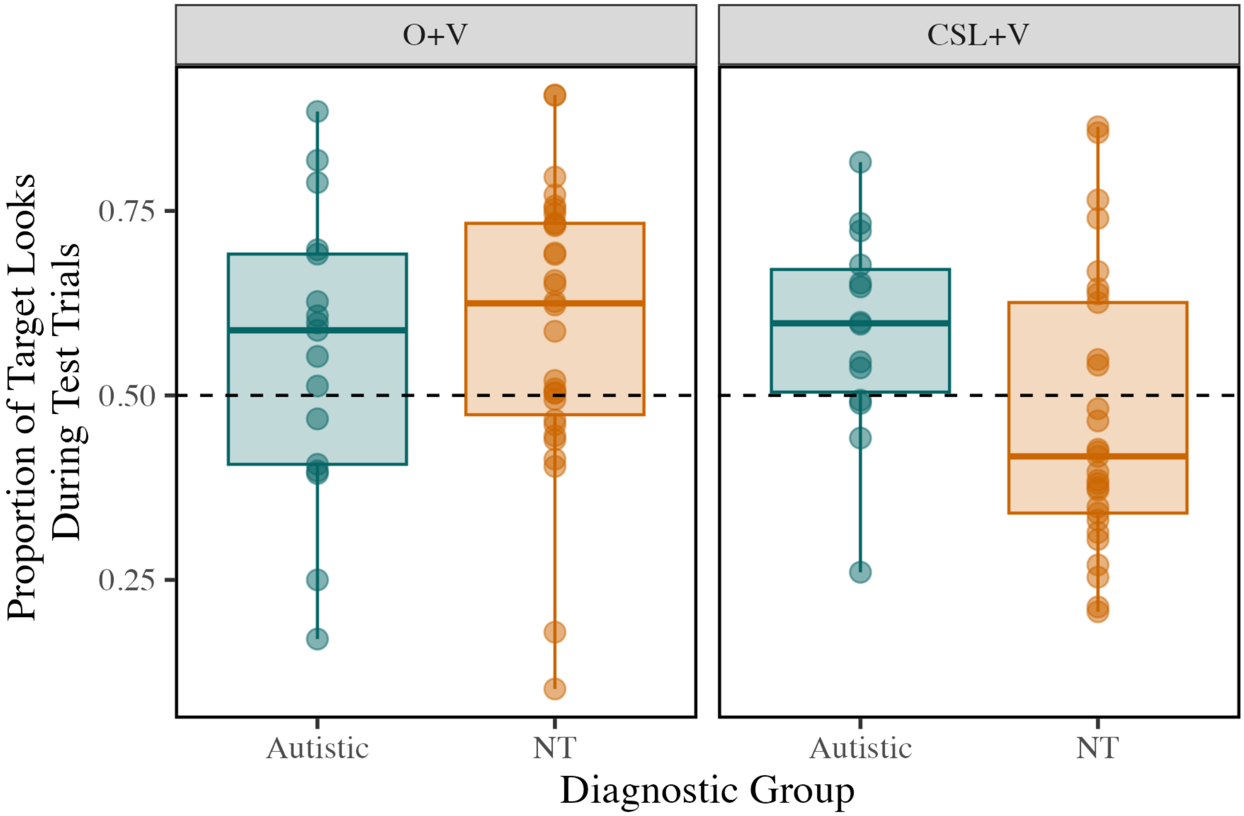


Figure 13. Proportion looks to target images averaged across the analysis window (300-1800 ms after target noun onset) during trained (i.e., fast mapping) test trials for NT (green) and autistic children (orange) with data included in analyses for the CSL+V task (left) and the O+V task (right). Each dot represents one participant's average in that task.

Chapter 5: Individual Child Characteristics Associated with Autistic Children's Word Learning

The purpose of this final study was to investigate child characteristics predictive of autistic children's performance in the three eye-gaze tasks comprising the previous studies to better understand within-group individual differences in word learning under conditions of referential ambiguity and perceptual variability. The three tasks were described in detail in General Methods (Chapter 2). The first task followed a classic cross-situational word learning (CSL) paradigm where two novel referents appeared side-by-side in each trial were labeled ambiguously with no cues as to which word labeled each referent. As such, children must track the statistical co-occurrences between novel words and referents across trials to form accurate mappings. This task will be referred to throughout this chapter as CSL-NV (CSL-no variability). The second task also followed a CSL paradigm and was identically structured to the CSL-NV task with one critical difference: throughout training, each novel word was paired with three different exemplar images that shared the same shape but differed in color. This task will be referred to as CSL+V (CSL with variability). The final task also featured exemplar variability, with three exemplar images varying in color but sharing the same shape paired with each novel word throughout training. However, this task presented only one word-referent pairing at a time during training. As such, the effect of exemplar variability could be isolated from the effect of referential ambiguity (i.e., CSL). This task will be referred to as O+V (ostensive teaching with variability). In sum, the three tasks could be conceptualized as representing word learning under conditions of referential ambiguity (CSL-NV), perceptual variability (O+V), and referential ambiguity with perceptual variability (CSL+V). Children's average accuracy on these tasks were visualized in plots in Chapter 3 and 4. The child characteristics included as predictors in analyses were receptive language (as measured by Preschool Language Scales – Fifth Edition (PLS-5; Zimmerman et al., 2011) Auditory Comprehension (AC) raw scores), level of autism traits (as measured by Childhood Autism Rating Scale – 2nd Edition (CARS-2); Schopler et al., 2010), nonverbal cognitive ability (as measured by Developmental Assessment of Young Children – 2nd Edition (DAYC-2; Voress & Maddox, 2013) cognitive domain raw scores), and chronological age (in months). Details of data collection for these standardized assessments can be found in General Methods (Chapter 2). I chose these characteristics because similar measures were investigated in previous studies of CSL in autistic children (Hartley et al., 2020; McGregor et al., 2013; Venker, 2019). These studies found that receptive vocabulary (Hartley et al., 2020; McGregor et al., 2013) and familiar word processing (Venker, 2019) predicted performance on CSL tasks at least in terms of fast mapping (McGregor et al., 2013; Venker, 2019) and referent selection during training (Hartley et al., 2020), though Hartley et al. (2020) found that retention and generalization of CSL were not related to any examined child characteristic. I therefore expected that receptive language would predict autistic children's performance in the CSL tasks. The effect of exemplar variability and its associated child characteristics have not yet been examined in autistic children, but one might expect level of autism traits to predict performance in the variability tasks if the mechanism underlying a lack of shape bias is specific to autism and also influences learning from variable exemplars as I had originally hypothesized.

I asked the following research questions:

1. Do receptive language, nonverbal cognitive ability, level of autism traits, or chronological age predict 3- to 4-year-old autistic children's performance over and above the other characteristics included in the model in a referentially ambiguous learning context (i.e., CSL-NV)?

2. Do receptive language, nonverbal cognitive ability, level of autism traits, or chronological age predict 3- to 4-year-old autistic children's performance over and above the other characteristics included in the model in a perceptually variable learning context (i.e., O+V)?
3. Do receptive language, nonverbal cognitive ability, level of autism traits, or chronological age predict 3- to 4-year-old autistic children's performance over and above the other characteristics included in the model in a referentially ambiguous *and* perceptually variable learning context (i.e., CSL+V)?

Methods

Participants

This study included data from 20 autistic children (5 female; m=44.05 months, SD=3.97, range=36-50). Description of recruitment procedures and exclusionary criteria can be found in General Methods (Chapter 2). It should be noted that the children included in the analyses for the CSL-NV task in this chapter were only the children who were in the Year 2 age range at the time of testing. I opted to include only those children to keep the sample as consistent as possible across the three research questions in the present study as the CSL+V and O+V tasks were only administered to children in the Year 2 age range of the larger longitudinal parent study. However, not all participants in the Year 2 age range ultimately had data for each study included in analyses due to data cleaning criteria (described below in Data Cleaning). After data cleaning criteria were applied, data were available for 17 autistic children for CSL-NV task, 14 autistic children for the CSL+V task, and 17 autistic children for the O+V task. Sample characteristics for participants with data included for the CSL-NV task can be found in Table 2. Sample characteristics for participants with data included for the CSL+V task and the O+V task can be found in Table 4 and Table 5.

Data Analysis

Data Cleaning

Data cleaning procedures were described in more detail in Chapter 3 and 4. To summarize: within each task, 33-ms frames during the analysis window (300-1800 ms after target noun onset during trained (i.e., fast-mapping) test trials) were removed from analyses if coders determined participants were shifting their gaze or looking away from visual stimuli. If a participant missed 50% or more frames during the analysis window of a test trial, that trial was excluded from analyses. Participants were excluded if they contributed 2 or fewer test trials (out of a total of 8) to increase the reliability of the average accuracy outcome measure. Average contributed trials after data cleaning by group and task were described in Chapter 4.

Analytic Approach

After data cleaning, looks to target (coded 1) and distractor (coded 0) from the analysis window (300-1800 ms after target noun onset) across all included trained (i.e., fast-mapping) test trials were averaged within each participant for each task. Thus, the outcome variable included one average accuracy measure (i.e., proportion looks to target) per participant for each task in which they contributed sufficient data for inclusion. For each of the three research questions, I fit a generalized linear model using the beta distribution and a logit link with the *glmmTMB* package in R (Brooks et al., 2017). Each model regressed average accuracy (in terms of log-odds of looking to target due to logit link) on receptive language (PLS-5 AC raw scores), nonverbal cognitive ability (DAYC-2 cognitive domain raw scores), level of autism traits (CARS-2 total scores), and age (in months). Each predictor variable was grand mean-centered and scaled on the full autistic sample. The models included outcome data from children with sufficient data for the CSL-NV task only for RQ1, the O+V task only for RQ2, and the CSL+V task only for RQ3. Unlike Venker (2019) who conducted independent correlations between accuracy and each child characteristic, I opted to examine all child characteristics together in one model. The reasoning for this analytical decision was based on the fact that these characteristics tend to intercorrelate, so I was interested in whether any specific child characteristic was related to word learning in the eye-gaze tasks over and above the others. The potential downside of this approach is that including intercorrelated predictors in the same model could result in multicollinearity, thus decreasing the reliability of the estimates and inference testing for the predictors (Thompson et al., 2017). To check for multicollinearity in the models, I measured the variance inflation factor (VIF) for each predictor. While there is no universally accepted threshold for VIF, >10 is frequently cited as an indication of high multicollinearity (e.g., Bayman & Dexter, 2021; Midi & Bagheri, 2010; Salmerón et al., 2020; Thompson et al., 2017) while a VIF between 5 and 10 may indicate moderate multicollinearity (Midi & Bagheri, 2010). However, Thompson et al. (2017) advise against using a measure like VIF dichotomously; that is, a score of 9.9 is not meaningfully different from a score of 10.1 even if a cutoff of 10 is applied. Across the three models fit in this study, the majority of predictors had VIF values <5, suggesting low multicollinearity. The only predictors with values above 5 were nonverbal cognitive ability (VIF=6.47) and receptive language (VIF=5.41) in the CSL+V model and nonverbal cognitive ability (VIF=5.18) in the O+V model. While these values were well below a theoretical threshold of 10, I conducted post-hoc exploratory analyses refitting the CSL+V and O+V models without each of the predictors with VIF above 5 to investigate whether model results would differ without their inclusion.

Results

For the first research question investigating child characteristics associated with autistic children's performance in the CSL-NV task, model results revealed a significant effect of receptive language, β=0.86, *SE*=0.42, z=2.07, p=.039. This finding indicates that autistic children with stronger receptive language skills demonstrated stronger performance in the CSL-NV task when accounting for nonverbal cognitive ability, level of autism traits, and chronological age. No other predictors were significant in this model (*p*s>.05). Full model results are in Table 8.

The second research question investigated child characteristics in relation to autistic children's performance in the O+V task. Model results indicated that chronological age significantly predicted task performance (β=0.40, *SE*=0.19, z=2.16, p=.031) when accounting for the effects of nonverbal cognitive ability, receptive language, and level of autism traits. No other predictors were significant (*p*s>.05). Full model results are in Table 9. Because nonverbal cognitive ability had a VIF of 5.18 indicating moderate multicollinearity, I re-fit the model excluding nonverbal cognitive ability. In this post-hoc model, all predictors had VIF values <2. Chronological age remained significant (β=0.39, *SE*=0.18, z=2.14, p=.033) when accounting for the effects of receptive language and level of autism traits.

The third research question examined the relationship of child characteristics with performance in the CSL+V task. The only significant predictor in this model was the intercept, β=0.36, *SE*=0.12, z=2.94, p=.003. Because the child characteristic variables were mean-centered, this indicates that a child theoretically at the average of the autistic sample in terms of nonverbal cognitive ability, receptive language, level of autism traits, and chronological age had log-odds of looking to target significantly greater than 0. None of the child characteristics were predictive of task performance (*p*s>.05). Full model results are in Table 10. Because two predictors in this model (nonverbal cognitive ability and receptive language) were moderately collinear (VIFs of 6.47 and 5.41 respectively), I refit the model two more times, first excluding nonverbal cognition and then including nonverbal cognition but excluding receptive language. In both models excluding one of the two collinear variables, all predictors had VIFs <2 indicating low multicollinearity. The pattern of results remained the same across all models; no child characteristics significantly predicted autistic children's task performance in the CSL+V task (ps>.05) regardless of whether nonverbal cognitive ability or receptive language were excluded.

Discussion

Utilizing the experimental eye-gaze data presented in Chapters 3 and 4, this study aimed to explore child characteristics predictive of individual differences in autistic children's word learning in various contexts. The eye-gaze tasks included input that was referentially ambiguous (i.e., the CSL-NV task), perceptually variable (i.e., the O+V task), and perceptually variable *and* referentially ambiguous (i.e., the CSL+V task). In doing so, I aimed to explain some of the heterogeneity in language abilities among young autistic children and begin to identify learning conditions most appropriate for specific autistic children based on their individual profile. Based on the child characteristics chosen for investigation in prior studies of CSL in autistic children (Hartley et al., 2020; McGregor et al., 2013; Venker, 2019), predictors of word learning task performance in this study were nonverbal cognitive ability, receptive language, level of autism traits, and chronological age. The first research question examined characteristics associated with performance in the CSL-NV task, the second research question examined characteristics associated with performance in the O+V task, and the third research question examined characteristics associated with performance in the CSL+V task. For each research question, a model was fit that included all four predictors. Thus, any significant findings suggest an effect over and above the other included characteristics.

As hypothesized, study results indicated that the performance among autistic children in the basic CSL task (CSL-NV) was predicted by receptive language when accounting for the effects of nonverbal cognitive ability, level of autism traits, and chronological age. This finding aligned with previous reports suggesting a relationship between aspects of receptive language ability and CSL in autistic school-age children and adolescents (Hartley et al., 2020; McGregor et al., 2013; Venker, 2019). For example, McGregor et al. (2013) reported a significant association between fast mapping accuracy in a CSL task and receptive vocabulary in autistic school-age children and adolescents (age in months: *M*=134, *SD*=25). Hartley et al. (2020) found that receptive vocabulary was predictive of referent selection accuracy during training in a CSL task among autistic school-age children, but not retention or generalization of learning. In Venker (2019), familiar word processing was correlated with accuracy in a CSL task among autistic school-age children. Venker (2019) reported a marginal but non-significant correlation of receptive vocabulary with CSL, which she attributed to the exclusion of six autistic children with low vocabulary scores limiting the variability in the receptive vocabulary measure. The measure of receptive language used in the present study (PLS-5 Auditory Comprehension raw scores) differed from prior work in that it encompasses broader linguistic skills than just vocabulary or familiar word processing. As such, it is not clear whether a more specific subskill (such as lexical processing or knowledge) might be driving the effect found in this study or whether multiple abilities within the domain might relate to CSL. However, the present findings suggest further investigation into the relationship between CSL and receptive language abilities beyond vocabulary in autistic children may be warranted. Aside from receptive language, no other child characteristics were associated with performance in the CSL-NV task. While null results are not directly interpretable, it is notable that prior CSL studies also did not find a significant effect of age (Hartley et al., 2020; McGregor et al., 2013; Venker, 2019), nonverbal cognitive ability (Hartley et al., 2020; Venker, 2019), or level of autism traits (Venker, 2019). Together with the extant literature, these findings point to an emerging relationship between receptive language and CSL superseding that of other characteristics.

The second research question investigated the same characteristics in relation to autistic children's performance in the O+V task, which involved explicit 1:1 labeling of perceptually variable exemplars during training. While no previous study has investigated the effect of exemplar variability on autistic children's word learning, I considered that level of autism traits might be relevant if the same mechanisms underlying a tendency not to extend labels based on shape (i.e., shape bias) also lead to difficulty learning from variable exemplars. However, as reported in Chapter 4, I did not find that autistic children had difficulty learning words in the context of exemplar variability. Thus, logic might follow that level of autism traits would not be predictive of children's performance. The only significant predictor that did emerge for the O+V task was chronological age (in months). The direction of this effect suggests that young autistic children become more successful at learning words from multiple perceptually variable exemplars with age - at least in a basic word learning task. Previous studies have demonstrated a similar developmental trend in NT children. For example, 10-month-old infants were able to map words to objects when trained on a single exemplar, but were not when presented with multiple variable exemplars during training (Taxitari et al., 2020). Two-year-old toddlers, on the other hand, mapped successfully in both conditions (Twomey et al., 2014). Vlach and Sandhofer (2011) found a similar pattern in a word learning task with variable background contexts. They found that 2.5- to 3-year-olds were detrimentally affected when the training context varied or the test context mismatched the training context, while 4-year-olds were successful across conditions.

The final research question sought to understand child characteristics related to autistic children's performance in the CSL+V task, which combined exemplar variability with a CSL paradigm, essentially combining the conditions of the two previous tasks. Unlike the CSL-NV and O+V tasks, I did not find evidence of any relationship between nonverbal cognitive ability, receptive language, level of autism traits, or chronological age with children's accuracy in the CSL+V task. Interestingly, this was also the only task in which the autistic children with sufficient data for inclusion in analyses were as a group significantly more accurate than chance. It may be argued that individual predictors of performance on tasks where the group average was not above chance represent random behavior. However, as noted in previous chapters, individual autistic children showed a wide range of accuracy, and several were individually above chance in absolute terms. I argue that it is just as important to investigate individual differences in these tasks for a few reasons: First, because we did not limit the sample to children who were able to learn in the tasks, the range of performance is representative of the heterogeneity of the autistic population (though the sample was limited in other ways; see previous chapters and General Discussion (Chapter 6). Second, these investigations may provide valuable insight into the characteristics that predict learning across the full range of task performance. Finally, examining individual differences among children with a wide range of accuracy allows for more robust statistical inferences. This final point may explain the null findings in the analyses of child characteristics related to the CSL+V task. While there was considerable within-group variability across all tasks (see Figure 12 and 13), the *SD* of children's average accuracy was 0.21 and 0.19

for the CSL-NV and O+V tasks, respectively, while the *SD* of the CSL+V task was 0.14. Thus, there may have simply been insufficient variance in the average accuracy measure for this task to detect any effect of individual child characteristics. It should be noted again that only 14 autistic children contributed sufficient data for inclusion in analyses for the CSL+V task compared to 17 each for the CSL-NV task and the O+V task. Therefore, the sample included may have also been slightly more homogenous than those who were included in the other two tasks. Of course, there may be some other characteristic shared by children who performed well in this task that was not captured among the included variables. Visual attention is a notable possibility given the differences in number of included children between tasks was based on amount of time spent looking to the visual stimuli. Assuming the CSL+V task was more perceptually salient than the other two tasks as discussed in Chapter 4, there may be within-group differences in learning depending on how strongly children's attention is pulled by stimulus salience. As discussed in Chapter 4, the novelty and salience of the exemplars presented during training was presumably relatively consistent throughout the trials and should therefore be less likely to compete with the unfolding statistical regularities for attention. As such, the novelty and salience of the exemplars may have acted instead as an attentional aid to the regularities informing the correct word-referent mappings particularly for children who are sensitive to low-level perceptual input. Given that some autistic children may have a variety of sensory profiles including both sensory-seeking and sensory-avoidant (Little et al., 2017), it is possible that children attended to and learned from the CSL+V task relative to their sensory profile.

Limitations and Future Directions

The findings presented in this study should be considered in the context of several limitations. First, it should be noted that the analytical decision to include all covariates in one model for each task resulted in a slightly different interpretation from the approaches of Venker (2019) or Hartley et al. (2020). Venker (2019) conducted individual correlations with each examined child variable and CSL task accuracy, while Hartley et al. (2020) took a model comparison approach iteratively adding child characteristic variables to the larger model and measuring improvement in model fit as an indication of individual differences in those characteristics. McGregor et al. (2013) regressed accuracy on age and receptive vocabulary both together and separately. A drawback of these approaches is the risk of inflating type-I error by conducting multiple comparisons, especially as it is unclear if the researchers conducted corrections on resulting *p*-values. However, the interpretation of the present findings is not precisely comparable to these previous studies as a result. Rather, all significant child characteristics found represent a relationship between that ability and task performance *when accounting for the effects of all the other included covariates*. This distinction is important as results may not have been identical if I had conducted individual correlations between each characteristic and child accuracy. Another limitation worth acknowledging is the differing sample sizes across tasks. Overall, included sample sizes were small (*n*=17 in the CSL-NV task, *n*=17 in the O+V task, and *n*=14 in the CSL+V task) but not very different from the autistic groups of similar studies (*n*=18, Venker, 2019; *n*=15, Hartley et al., 2020). It should be noted that I only included children in the analysis for the CSL-NV task who were tested at the Year 2 time point of the larger parent study, although additional children were tested at the Year 1 time point and were included in the main (matched groups) analysis in Chapter 3. This decision was due to the fact that children only participated in the other two tasks at the Year 2 time point and I sought to keep the included children in this study as consistent as possible across research questions. Even so, the samples were not equivalent across tasks due to different children meeting data cleaning criteria for each. Additionally, the predictor variables were centered and scaled on the full autistic sample, though some children were excluded in each analysis due to insufficient data for inclusion. While this analytical approach maintained consistency in the individual values for each predictor across models, the slight differences in sample characteristics due to exclusion may warrant individually centering and scaling the predictors for each model in future work. In sum, any qualitative comparison of the effects across the different models should be made cautiously as the analyses were not within-subjects, and while the centering of predictors was identical across models, the samples were not.

It should also be acknowledged that any causality or directionality of the relationship between the individual child characteristics and task performance cannot be gleaned from these analyses. Though one might interpret the findings as an indication that autistic children with stronger receptive language skills are better at tracking statistical regularities and extracting word-referent mappings under conditions of referential ambiguity (i.e., CSL), it could also be that some mechanism underlying CSL results in better receptive language ability as measured by standardized assessment (PLS-5 Auditory Comprehension raw scores in this case). Moreover, autistic children may learn better from exemplar variability with age over the preschool years (at least when referents are labeled explicitly, i.e., the O+V task), but it is not clear whether maturation alone is necessarily the cause of this improvement. It would be reasonable to say that it influenced performance over and above that of receptive language, nonverbal cognitive ability, and level of autism, but other unspecified factors that tend to change with development (e.g., attention) could still contribute.

Finally, while the present study provided some insight into individual differences associated with autistic children's initial word learning across conditions of referential ambiguity (i.e., CSL) and perceptual variability, it is important to acknowledge that learning a word involves considerably more than immediate referent selection. As discussed previously, an experimental design error prevented the planned investigation of children's generalization of learning in this dissertation, and therefore I cannot rule out the possibility that different child characteristics might emerge as predictive of performance on generalization test trials. In addition, future studies should examine individual differences associated with longer-term retention of learning. For example, Vlach and DeBrock (2019) demonstrated that NT children do not show reliable retention of words acquired through CSL after a 5-minute delay until 4 years old. However, in a study with autistic children, Hartley et al. (2020) found that receptive vocabulary predicted referent selection accuracy during training trials, but neither receptive vocabulary, nonverbal cognition, nor chronological age predicted accuracy in retention and generalization test trials.

Finally, the present results may aid speech-language pathologists in better understanding which word learning conditions might be most facilitative for young autistic children based on their individual ability profile. Further, intervention studies tailored to individual profiles would help determine causality and directionality in these relationships. Study findings suggest that autistic children with stronger receptive language skills are better able to extract statistical regularities from referentially ambiguous input. Future intervention designs might consider capitalizing on these children's strengths in detecting patterns in language input. For example, Alt et al. (2014) found a language intervention based on CSL principles resulted in vocabulary gains for late talkers. The present findings suggest that at least some autistic children might benefit from such an intervention design as well. Autistic children also appeared to improve with age in the ability to learn words from input involving variable exemplars (when they were explicitly named, i.e., in the O+V task). These results may support including more complex sets of within-lexical category stimuli with autistic children as they get older. Clinical implications of this work will be discussed further in General Discussion (Chapter 6).

Conclusions

Study results indicated that young autistic children with stronger receptive language abilities were better able to learn words from CSL without exemplar variability when accounting for the effects of nonverbal cognitive ability, level of autism traits, and age. Children were also found to be more successful in learning words from explicitly labeled variable exemplars with age when accounting for the effects of receptive language, nonverbal cognitive ability, and level of autism traits. These findings may have particular relevance for language interventionists in identifying input qualities most suitable for individual autistic children according to their unique profile. More research will be needed to confirm these findings in the broader population, determine direction and causality, and investigate other characteristics that might explain child performance in the CSL task with exemplar variability, such as sensory processing or visual attention.

References

Alt, M., Meyers, C., Oglivie, T., Nicholas, K., & Arizmendi, G. (2014). Cross-situational statistically based word learning intervention for late-talking toddlers. *Journal of Communication Disorders*, *52*, 207–220. https://doi.org/10.1016/j.jcomdis.2014.07.002

Bayman, E. O., & Dexter, F. (2021). Multicollinearity in Logistic Regression Models. *Anesthesia & Analgesia*, *133*(2), 362–365. https://doi.org/10.1213/ANE.0000000000005593

Brooks, M., E., Kristensen, K., Benthem, K., J. ,van, Magnusson, A., Berg, C., W., Nielsen, A., Skaug, H., J., Mächler, M., & Bolker, B., M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378. https://doi.org/10.32614/RJ-2017-066

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*(July 2019), 104265. https://doi.org/10.1016/j.cognition.2020.104265

Little, L. M., Dean, E., Tomchek, S. D., & Dunn, W. (2017). Classifying sensory profiles of children in the general population. *Child: Care, Health and Development*, *43*(1), 81–88. https://doi.org/10.1111/cch.12391

McGregor, K. K., Rost, G., Arenas, R., Farris-Trimble, A., & Stiles, D. (2013). Children with ASD can use gaze in support of word recognition and learning. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(7), 745–753. https://doi.org/10.1111/jcpp.12073

Midi, H., & Bagheri, A. (2010). *Robust Multicollinearity Diagnostic Measure in Collinear Data set*. 138–142.

Salmerón, R., García, C., & García, J. (2020). *Overcoming the inconsistences of the variance inflation factor: A redefined VIF and a test to detect statistical troubling multicollinearity* (arXiv:2005.02245). arXiv. http://arxiv.org/abs/2005.02245

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale—Second Edition (CARS-2)*. Western Psychological Services.

Taxitari, L., Twomey, K. E., Westermann, G., & Mani, N. (2020). The Limits of Infants' Early Word Learning. *Language Learning and Development*, *16*(1), 1–21. https://doi.org/10.1080/15475441.2019.1670184

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the Variance Inflation Factor and Other Multicollinearity Diagnostics from Typical Regression Results. *Basic and Applied Social Psychology*, *39*(2), 81–90. https://doi.org/10.1080/01973533.2016.1277529

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's More Like It: Multiple Exemplars Facilitate Word Learning. *Infant and Child Development*, *23*, 105–122. https://doi.org/10.1002/icd

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*(November 2018), 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 700–711. https://doi.org/10.1037/xlm0000611.Statistics

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Voress, J. K., & Maddox, T. (2013). *Developmental Assessment of Young Children-Second Edition (DAYC-2)*. PRO-ED.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Tables and Figures

Table 8. Full model results for model investigating child characteristics predictive of autistic children's performance in the CSL-NV task.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.22	0.19	1.14	.253
Nonverbal Cognitive Ability	-0.69	0.42	-1.67	.095
Receptive Language	0.86	0.42	2.07	**.039**
Level of Autism Traits	0.10	0.23	0.45	.655
Chronological Age	0.23	0.19	1.18	.237

Dispersion Parameter: 5.79

Table 9. Full model results for model investigating child characteristics predictive of autistic children's performance in the O+V task.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.17	0.16	1.06	.291
Nonverbal Cognitive Ability	0.52	0.39	1.33	.185
Receptive Language	-0.31	0.34	-0.89	.372
Level of Autism Traits	0.12	0.24	0.52	.600
Chronological Age	0.40	0.19	2.16	**.031**

Dispersion Parameter: 9.12

Table 10. Full model results for model investigating child characteristics predictive of autistic children's performance in the CSL+V task.

	Log-Odds of Target Looking			
	Estimates	*SE*	*Statistic*	*p*
Intercept	0.36	0.12	2.94	**.003**
Nonverbal Cognitive Ability	0.16	0.30	0.54	.587
Receptive Language	-0.01	0.28	-0.04	.967
Level of Autism Traits	-0.05	0.14	-0.36	.716
Chronological Age	0.23	0.14	1.58	.115

Dispersion Parameter: 19.5

Chapter 6: General Discussion

Through multiple eye-gaze experiments and standardized assessments, the studies comprising this dissertation investigated whether differences in underlying mechanisms might contribute to word learning among autistic children depending on characteristics of the learning environment and individual differences. As outlined in Chapter 1, these studies addressed this broader question within a framework informed by statistical learning theory (e.g., Saffran et al., 1996) and the Attentional Learning Account (Smith, 1999; Smith & Samuelson, 2006). In the first study, I investigated children's ability to extract statistical regularities from input when novel referents are labeled ambiguously across trials to form accurate mappings, referred to as cross-situational learning (CSL). I examined CSL in a younger cohort of autistic children than previous studies and compared their learning to younger neurotypical (NT) peers matching on or statistically accounting for nonverbal cognitive ability. Given developmental effects found in CSL studies with NT children of this age (Akhtar & Montague, 1999; Vlach & DeBrock, 2019) and the prevalence of early language delays in autism (Howlin, 2003; Mayo et al., 2013), I hypothesized that autistic children may have relatively more difficulty learning words from CSL. The second study investigated the impact of perceptually variable exemplars during training in both a CSL task and an ostensive teaching task where referents were explicitly labeled. I expected autistic children to demonstrate more difficulty learning words from variable exemplars compared to multiple exposures to a single exemplar and more difficulty learning from variable exemplars in comparison to NT peers. I also anticipated that autistic children would demonstrate an improvement in performance when referents were labeled explicitly compared to ambiguously (i.e., CSL). In the final study, I analyzed the relationship of child characteristics to individual differences in word learning from CSL and variable exemplars among young autistic children. Following prior studies, I investigated the effects of receptive language, nonverbal cognitive ability, level of autism traits, and chronological age on autistic children's performance in the three experimental eye-gaze tasks reported in the previous two studies of this dissertation. I hypothesized that receptive language would predict children's performance in the CSL tasks and that level of autism traits might predict performance in the tasks that included exemplar variability.

Results of the first study revealed that 2- to 4-year-old autistic children did not show evidence of learning from cross-situational statistics as measured by accuracy during fast-mapping (i.e., referent selection) test trials (defined as proportion looking to target images). However, NT children who were matched on nonverbal cognitive ability also did not show evidence of learning and did not differ from autistic peers. In a secondary analysis, I examined CSL among all children within the age range of the Year 2 time point of the larger project under the assumption that older children would demonstrate stronger performance in the task. Indeed, NT children in the age range of the Year 2 time point (30-39 months) did demonstrate above-chance average accuracy, indicating that NT children in this age range were able to learn words from this CSL task. By contrast, the autistic children tested at the Year 2 time point (when they were 36-50 months) did not as a group differ from chance in their average accuracy. However, the autistic children did show considerable variability in performance as visualized in Figure 9. Moreover, they did not as a group differ from NT peers when nonverbal cognitive ability was included as a covariate. These findings suggest that many autistic children may have difficulty learning words from cross-situational statistics before the age of school entry, at least as defined by this paradigm. However, several children individually demonstrated accuracy that exceeded chance (.50) in absolute terms, indicating that such a conclusion would not apply universally across the autism spectrum.

Findings of the second study were more surprising. Contrary to expectations, 3- to 4-year-old autistic children were able to learn words in a CSL task that included perceptually variable exemplars during training (CSL+V) as demonstrated by average accuracy significantly above chance. The NT children, on the other hand, did not differ from chance in their average accuracy. In fact, the NT children showed the opposite pattern in accuracy between tasks. That is, the NT children tested when in the age range of the Year 2 time point were above chance in average accuracy in the CSL task without variability (CSL-NV) but did not differ from chance in the CSL+V task. The autistic children also outperformed the younger NT children in the CSL+V task when covarying nonverbal cognitive ability and accounting for other model terms. Autistic children did not show a difference in performance between the CSL-NV and the CSL+V task, however, likely due to the wide variability in the CSL-NV task. Also contrary to my original hypotheses, the autistic children did not show an improvement in word learning when variable exemplars were labeled explicitly (O+V task) compared to ambiguously (i.e., CSL+V task). Surprisingly, the autistic children did not differ from chance in average accuracy in the O+V task, though the NT children did. The NT children also showed more of a benefit from explicit labeling in comparison to the autistic group when accounting for all other model terms and covarying nonverbal cognitive ability.

The third study revealed partial support for my original hypotheses. That is, receptive language predicted autistic children's performance in the CSL-NV task when accounting for the effects of nonverbal cognitive ability, level of autism traits, and chronological age. However, neither receptive language nor any other characteristic examined significantly predicted performance in the CSL+V task. For the O+V task, chronological age predicted autistic children's performance such that older children showed stronger learning than younger children. I had anticipated that level of autism traits might influence children's accuracy on the tasks that included exemplar variability under the assumption that the potential mechanisms underlying differences in shape bias task performance such as those described by Tek and Naigles (2017) or Tovar et al. (2020) might also lead to difficulty learning from perceptually variable exemplars. However, given that autistic children did not demonstrate the expected pattern in relation to exemplar variability, it might follow that level of autism traits would not be relevant.

Some results of this dissertation replicate previous work, while others contribute novel findings that open doors to future lines of research. Like prior studies (e.g., Smith & Yu, 2008; Venker, 2019; Vlach & DeBrock, 2017), I found that young NT children are able to use cross-situational statistics to learn words. However, the experimental design of these studies appears to have been difficult for the youngest NT children as group-level average accuracy did not exceed chance in the cognitive ability-matched group which included children as young as 19 months with an average age of 27.63 months. It was only when I examined NT children who were within the age range of the Year 2 time point in the parent study (with an average age of 32.96 months) that average accuracy was greater than chance in the basic CSL task (CSL-NV). As prior work reported evidence of learning in same-age or even younger NT children (Smith & Yu, 2008; Vlach & DeBrock, 2017; Vlach & Johnson, 2013), the present findings should not be interpreted as suggesting that the youngest NT children in this study were not capable of CSL. Rather, researchers should be aware that CSL is sensitive to differences in methodology and analytical approach, particularly among children in the early stages of language development. Another finding of this dissertation confirming extant evidence was the relationship of CSL (without variability) to receptive language abilities within the autistic group. For example, McGregor et al. (2013) found that receptive vocabulary predicted fast-mapping accuracy in a CSL task among school-age children and adolescents, while Hartley et al. (2020) reported a relationship between receptive vocabulary and referent selection accuracy during training in a CSL task among school-age children. Venker (2019) found only a marginal, non-significant relationship between receptive vocabulary and CSL task performance in school-age autistic children, but a significant relationship of CSL task performance with familiar word processing. This dissertation extended this emerging body of evidence in several ways: first, whereas previous studies connected CSL to only receptive vocabulary or familiar word processing, the present work demonstrated that CSL was related to receptive language skills broadly (as measured by Preschool Language Scales 5th Edition (PLS-5; Zimmerman et al., 2011) Auditory Comprehension raw scores). Moreover, results indicated that this relationship was significant when accounting for nonverbal cognitive ability, level of autism traits, and chronological age. In addition, these findings demonstrated that CSL was related to receptive language in 3- to 4-year-old autistic children – a younger cohort than any previous study. Together, this growing body of evidence suggests an important role for CSL in the receptive language development of autistic children, which could be explored in more depth in future work including investigation of the directionality of this effect.

Additionally, this dissertation contributed novel findings informing contexts supportive of word learning in autism. The primary novel contribution of this work was the above-chance performance of autistic children in the CSL task that included exemplar variability (CSL+V) which was superior in relation to the NT children when covarying nonverbal cognitive ability and accounting for all other model terms (note however that the NT children were younger than the autistic children – whether this pattern would hold with age-matched NT peers remains an open question). While I did not find a significant difference in autistic children's learning between the CSL+V task and either the CSL-NV or O+V tasks, it was notable that their accuracy exceeded chance in the CSL+V task but not the CSL-NV or O+V task. By contrast, NT children in the age range of the Year 2 time point were significantly more accurate than chance in the CSL-NV and O+V tasks but not the CSL+V task. These findings were surprising; I had expected exemplar variability to negatively affect autistic children's word learning compared to no variability and to a greater extent than NT children given previous literature reporting a lack of shape bias at the group level in this population possibly attributed to underlying conceptual or processing differences. For example, Tek and Naigles (2017) suggest that difficulties associated with lexical organization such as abstracting category-relevant properties across exemplars may be responsible. Relatedly, computational modeling of shape bias differences indicated that autistic individuals may have a processing style characterized by hyperplasticity and hyperconnectivity leading to overly-specified categorical representations that compete to process input (Tovar et al., 2020) preventing integration of exemplars into an abstracted categorical prototype. I hypothesized that if these researchers' mechanistic explanations for the shape bias effect are correct, then by extension learning a word associated with a set of perceptually variable within-lexical category exemplars may be more difficult for young autistic children in comparison to learning a word associated with a single exemplar. However, these data did not support this hypothesis. Not only were autistic children able to learn words via cross-situational statistics with exemplar variability, but at least among children with sufficient data for inclusion, they did so at a higher level of accuracy than NT peers when covarying nonverbal cognitive ability. Why might autistic children have been more successful in this task than anticipated? One possibility is they relied on different mechanisms to learn words in the CSL+V task than children in shape bias studies. Because of the experimental design error in the planned generalization trials in the present studies, I was only able to analyze fast mapping test trials. In these trials, children only had to show recognition of an association of the target label and one exemplar that had appeared during training over a distractor that had been given a different label. By contrast, shape bias study test trials pair a novel exemplar that shared the same shape as the training exemplar with an object that shared some other characteristic with the trained exemplar, such as color. Thus, children may have been able to rely more on statistical, associative learning to succeed in the fast mapping trials of the CSL+V task in comparison to shape bias studies. The training exemplars in the CSL+V task also varied within-category relatively narrowly, differing only in color (Appendix A). Perhaps varying on more dimensions would have pulled more attention or processing load and elicited the expected effect. As it was, the combination of narrow exemplar variability and the CSL paradigm involving two images on screen at a time both during training and test trials resulted in learning in the autistic group, at least among those with sufficient data for inclusion in analyses. In Chapter 4, I speculated that differences in visual attention to perceptually salient stimuli may have contributed to the superior performance of autistic children in the CSL+V task. This possibility has mixed support from previous studies exploring the role of perceptual salience in linguistic processing and word learning with autistic children (Venker et al., 2021, 2022). In Venker et al. (2021), perceptually salient familiar object images attracted attention from named (but less salient) familiar object images in autistic children to a greater degree than NT children. However, Venker et al. (2022) explored the role of perceptual salience in autistic children’s novel word learning and found that high perceptual salience was detrimental compared to low perceptual salience, which the authors attributed to “sticky” attention to high salience distractor images. As I discussed in Chapter 4, differences in experimental design may have prevented an effect of “sticky” attention in the CSL+V task, likely contributing to the discrepancy in findings compared to those reported by Venker et al. (2022). Unlike other studies where novelty/salience in the training set attracted children’s attention to the detriment of their word learning (Smith & Yu, 2013; Venker et al., 2022), the salience and visual complexity of the CSL+V task were relatively consistent through both training and test phases. Moreover, the CSL+V task included more exposures to each word-object pairing than Venker et al. (2022). As such, the design afforded children with ample opportunities to redirect their attention across stimuli during training and become habituated to the level of novelty and visual complexity by the test phase.

As I noted in Chapter 4, it is important to interpret these results with consideration for the effect of the data cleaning criteria applied. I opted to include children’s data in analyses at the task level only if they contributed more than 2 (out of 8) usable trials (that is, trials in which they were looking at visual stimuli for at least 50% of the analysis window). This decision led to unequal sample sizes across tasks as some children met criteria for some tasks but not others. Of note, only 14 autistic children met criteria for inclusion in analyses for the CSL+V task while 17 met criteria for the CSL-NV and O+V tasks. Thus, one possible alternative explanation for the above-chance accuracy among autistic children in the CSL+V task is that the lowest-performing children in the other two tasks simply looked away from both target and distractor images more often instead of looking more equally between the two (or looking at the distractor more often) as they did in the other tasks. To explore this possibility, I plotted the average accuracy scores of children who contributed sufficient data for inclusion in spaghetti plots to visualize the within-child differences across studies. In these plots, children’s average accuracies within a task were indicated with a dot, while a line connected the average accuracies in each task for a given child. If a child only contributed data for one task, their average accuracy would be denoted by a single dot without a line. By examining where the single dots without connecting lines lie on the plot, one can visualize the average accuracy of children who contributed for only the CSL-NV but not the CSL+V task (Figure 14) or the O+V task but not the CSL+V task (Figure 15). In Figure 14, it does not appear as though autistic children who contributed data for the CSL-NV task only were qualitatively less accurate than the rest of the sample. Again speaking qualitatively, Figure 15 suggests that at least one of the autistic children who contributed data for the O+V task but not the CSL+V task was on the lower end of the overall spread in average accuracy for the sample, while the others were above chance (.50) and more accurate than about a third of the rest of the sample. Overall, these plots do not appear to suggest that only the children who were less accurate in the other tasks were excluded from the CSL+V task. Rather, while I speculated that autistic children may have been more accurate in the CSL+V task at least in part because its increased salience promoted visual attention, autistic children can present with a variety of sensory profiles. Some children may have a greater tendency to seek out sensory input, others may have a stronger tendency to avoid sensory input, and others may present with a mix of both (Little et al., 2017). As discussed in Chapter 4, it is possible that children who looked away from visual stimuli most (and were therefore excluded from analyses) in the CSL+V task may have found it to be an overwhelming sensory experience. These children's inattention may also be related to other skill domains. In a basic CSL eye-gaze experiment with a very similar structure, Venker (2019) reported that autistic children who were excluded from analyses due to insufficient data contributed after cleaning had lower language skills. While not the intended focus of this dissertation, findings may warrant future research to better understand the role of perceptual salience in autistic children's visual attention and its connection to word learning – particularly with consideration for individual differences in sensory processing.

A secondary novel finding of this research was that naming perceptually variable objects explicitly (as opposed to ambiguously, i.e., CSL) improved learning for NT children to a greater extent than autistic children. That NT children demonstrated learning in both a CSL task *without* exemplar variability and an ostensive teaching task *with* exemplar variability but not a combination of the two may suggest an effect of cognitive load due to the complexity of the CSL+V task for the NT children. Why might this pattern differ for autistic children? I outlined a few possibilities in Chapter 4, all related to attentional differences. First, the order of administration of the tasks was such that half the children saw the O+V task first and half saw it last. Some children may therefore have been fatigued by the time they saw that task, and groups may have differed in the extent to which that fatigue affected their attention given previous findings of diagnostic group differences in sustained attention (Chien et al., 2014, 2015; Vivanti et al., 2017; but see Johnson et al., 2007) particularly given the prevalence of co-occurring attention deficit/hyperactivity disorder (ADHD) in autism (Yerys et al., 2009). Another possibility is that changing from one object on screen during training to two objects on screen during test in the O+V task invited more visual exploration of both images or "sticky" attention as was found in Venker et al. (2022).

A final novel finding of this dissertation was that age predicted learning from explicit naming of variable exemplars (i.e., the O+V task) in autistic children when accounting for nonverbal cognitive ability, receptive language, and level of autism traits. This result may suggest that autistic children are better able to learn words from variable exemplars as they get older, and that this relationship is not better explained by developmental change in receptive language, nonverbal cognitive ability, or level of autism traits. There may be other unmeasured developmental factors that account for this relationship such as improved attention or executive function, which could be explored further in future studies.

Finally, the studies comprising this dissertation yielded several null results that – while not directly interpretable – are worth noting for future research. First, neither of the cognitive ability-matched groups in the CSL-NV task (reported in Chapter 3) differed from chance in average accuracy as a group, nor did the autistic children included in the secondary analysis. Further, while the NT group in the secondary analysis did show evidence of learning, their performance in the CSL-NV task did not significantly differ from autistic peers when covarying nonverbal cognitive ability. While it is possible that the pattern of results in this study together with Venker (2019) may suggest CSL (as measured by this paradigm) comes online at an age averaging somewhere between the present study (range=27-50 months) and that of Venker (2019; *m*=76 months, range=48-97 months), it is equally as possible that a simpler task, more repetitions, or a longer test trial duration might have facilitated successful learning in the autistic group. Future researchers should be cautious in designing CSL experiments for young children as methodological differences between studies can clearly have a significant impact on results. As discussed above, the autistic group also did not differ from chance in the O+V task and did not improve in performance compared to the CSL+V task. These null results could be attributed to the explanations in the previous paragraph, but in any case appear to apply to only a few children. Visualizations of the average accuracy data (Figure 12 and 13) suggest that the low scores of a small number of autistic children brought down the average for the O+V task. It is worth noting again that the median average accuracies for autistic children were very similar between tasks (*Mdn*=.60 in CSL+V and *Mdn*=.59 in O+V). A final null finding that should be highlighted was that no included child characteristic (nonverbal cognitive ability, receptive language, chronological age, or level of autism traits) predicted autistic children's performance in the CSL+V task when accounting for the effects of the other covariates. The lack of findings in that analysis could indicate that other unmeasured variables such as attention or sensory processing are more relevant to performance in this task. Alternatively, there may not have been sufficient variance in the outcome measure to detect relationships with child characteristics. Future research may improve upon this work by including larger sample with a wider range of task performance. As a final note, level of autism traits and nonverbal cognitive ability were not predictive of performance in any of the word learning tasks when accounting for the other covariates. It is unclear whether these characteristics were truly unrelated to word learning in these tasks or if some other factor prevented detection of significant effects. For example, the measures used to operationalize these constructs were both chosen for their reliability in remote data collection necessitated by the COVID-19 pandemic. However, neither measure has historically been used by our research group and may not be capturing these abilities in the same way as more commonly used measures that are administered in-person.

Theoretical Contributions

Results of these studies provide partial support for an extension of statistical learning theory to young autistic children. That is, in these experiments autistic children as a group did show evidence of word learning based on statistical co-occurrences among words and referents in the input, but only when exemplars varied in color during training. However, within-group variability in performance was evident across tasks, suggesting that CSL may be an available word learning process to some (but not all) autistic children. Further, the significant relationship of receptive language with performance in the basic CSL task (when accounting for other covariates) could suggest that autistic children's ability to use cross-situational statistics results in stronger language skills, though the current study cannot confirm causality. Previous findings offer some additional support for this view, relating various aspects of receptive language and CSL task performance in autistic children such as receptive vocabulary (Hartley et al., 2020; McGregor et al., 2013) and familiar word processing (Venker, 2019). Venker (2019) found familiar word processing to also predict CSL in NT children, but receptive vocabulary was not predictive of CSL in NT children across studies (Hartley et al., 2020; McGregor et al., 2013; Venker, 2019). Additionally, children with DLD have demonstrated above-chance learning on CSL tasks but weaker performance relative to peers with typical development (Ahufinger et al., 2021; Broedelet et al., 2023; McGregor et al., 2022). Together, these findings indicate that cross-situational statistics are a viable cue for word learning across diagnostic groups and may be related to broader language abilities in some children.

Unfortunately, the error on the generalization test trials precluded these studies from providing much insight into the feasibility of the Attentional Learning Account (ALA; e.g., Smith & Samuelson, 2006) for young autistic children. However, results may suggest that attention itself could play a role in autistic children's word learning if, as I speculated in Chapter 4, exemplar variability increased the overall perceptual salience of the CSL+V task and improved attention to the task among autistic children as a result. Further research that includes generalization of learning from variable exemplars will provide a better indicator as to whether learning to attend to lexical category-relevant features across input enables word learning in autistic children.

Finally, these results do not support a connection between mechanisms underlying difficulties with shape bias tasks with learning from variable exemplars in autistic children as I had hypothesized. If computational modeling by Tovar et al. (2020) was correct in characterizing autistic lexical processing by fragmented, overly-specified representations competing to process incoming stimuli, then we might expect autistic children to process each incoming exemplar as a new lexical category rather than a member of an existing lexical category along with previous exemplars given the same label, preventing the formation of generalizable lexical representations and possibly even impeding immediate word-referent mapping. However, current results do not support this hypothesis at least as it relates to fast-mapping of variable exemplars. Again, future research that includes usable generalization test trials will provide stronger evidence for this claim.

Clinical Implications

The results of these studies may inform future clinical interventions for autistic children with co-occurring language difficulties. First, findings indicate that at least some autistic children were able to make use of the statistical regularities in input to correctly map novel words and referents. However, as a group they showed above-chance accuracy only when training exemplars varied perceptually, though across tasks many children had individual average accuracies that were above chance in absolute terms. Together, these findings suggest that interventions could be designed to harness strengths in CSL for some autistic preschool children. As discussed in Chapter 5, an intervention incorporating CSL has previously been investigated for late talkers which led to increases in vocabulary (Alt et al., 2014). Other autistic children clearly struggled with the referential ambiguity of the CSL tasks and may benefit more from other interventions. Clinicians and intervention researchers should also consider the perceptual salience of therapy stimuli given the above-chance learning among autistic children in the CSL task that included exemplar variability. Depending on their individual sensory profile, autistic children’s visual attention might be differentially affected by the level and consistency of the perceptual salience in the input which could in turn impact their word learning. However, this possibility should be investigated further in future research as this dissertation was not designed to investigate these constructs. Finally, the individual differences predictive of word learning reported in Chapter 5 may assist clinicians in determining input characteristics most facilitative for word learning for individual autistic children. For example, autistic children with stronger receptive language skills showed stronger word learning from cross-situational statistics (when exemplars did not vary). Thus, an intervention designed to incorporate CSL may be beneficial for autistic children with stronger baseline receptive language skills, while children with greater needs in receptive language may benefit from different approaches including more explicit naming or possibly more repetitions, longer exposures, additional cueing, or attentional support. While the above-chance learning in the autistic group for the CSL+V task overall would suggest that complex within-lexical category therapeutic stimuli sets could be appropriate for many children, older autistic children in the sample showed stronger word learning in the task that included variable exemplars that were explicitly named. Therefore, clinicians might consider increasing the complexity of therapeutic stimuli sets as children get older, or perhaps beginning with a single exemplar or a very simple set of stimuli with younger children if they have difficulty learning a target word. Further research will be necessary to determine whether implementing an individualized therapeutic approach in this manner would result in measurable gains, which would additionally help establish causality in the relationships detected in the Chapter 5 analyses.

Limitations and Future Directions

Several limitations should be considered when interpreting the results of this dissertation, several of which could be addressed in future work. First, it is critical to acknowledge that the demographic composition of the sample limited the generalizability of these findings to the broader population. The sample was primarily white and non-Hispanic or Latino from higher socioeconomic backgrounds as measured by maternal education (see Table 1, Table 2, Table 4, Table 5), and thus these findings may not apply to other demographic groups. Future research with more representative samples will be important not only for this work but in research on language development in autism more broadly (Girolamo et al., 2023). Sample sizes were also relatively small, particularly for the autistic group, which may have limited statistical power to detect significant effects. It is also important to reiterate that included children varied somewhat across tasks due primarily to data cleaning procedures. After first cleaning at the trial level (removing trials in which children were looking to visual stimuli for 50% or less of the analysis window), I chose a criterion of >25% for trial contribution at the participant level for inclusion in analyses such that children who contributed 2 or fewer trials (out of a total of 8 fast mapping trials) were excluded. I chose this criterion to increase the reliability of the average accuracy outcome measure, assuming that data from only one or two trials would be too noisy. However, a consequence of this analytical decision was that groups were not equivalent across tasks as some children contributed sufficient data for some tasks but not others. For the purpose of comparing performance across tasks, I analyzed all the available data with mixed-effects models assuming missing at random (MAR; Pugh et al., 2021) because the data were primarily missing due to a feature of the experimental design, one missing task was due to experimenter error, and two missing tasks were because the child was too fussy to continue. However, it cannot be entirely ruled out that children with excluded data for one task may have differed in some unmeasured systematic way that would violate the MAR assumption. Additionally, the predictor variables in Chapter 5 were centered and scaled based on the entire sample, though again the models for each study had some missing participants due to insufficient data. Results of all studies should be considered cautiously with this possible bias in mind. This caution should be applied particularly to the CSL+V task as the fewest autistic children met criteria for inclusion in this task and it was the only task in which autistic children as a group showed above-chance accuracy.

Another important limitation of these studies which I have discussed in previous chapters is the lack of matching between diagnostic groups on any child characteristic except for the primary analysis reported in Chapter 3. Children who participated in these studies were primarily recruited through a larger longitudinal parent study which was designed to match diagnostic groups on nonverbal cognition. Thus, children in each group were recruited within an age range in which their nonverbal cognitive ability would be approximately matchable. This design choice resulted in a significantly younger NT sample compared to the autistic sample. Despite falling within a closer cognitive ability range, the groups remained significantly different on all measured variables. Moreover, only children in the Year 2 age range of the larger parent project participated in the CSL+V and O+V task, limiting the sample size even further. As such, matching groups was not feasible in terms of retaining adequate statistical power. I opted instead to include all children with sufficient data for analysis and include nonverbal cognitive ability as a covariate in the models. While this approach could be considered an alternative method of accounting for group differences in nonverbal cognitive ability, researchers have argued that variables such as cognitive ability are not meaningfully separable from diagnostic group thus complicating interpretation of results (Dennis et al., 2009; Miller & Chapman, 2001). I also did not include any other child characteristics as covariates in the group comparisons since these variables were so confounded with diagnostic group that their inclusion would likely have absorbed too much variance to detect any effect of diagnostic group or contributed to multicollinearity in the models. Thus, diagnostic group effects reported in these studies should be considered with regard to differences on the variables reported in Table 1, Table 2, Table 4, and Table 5.

Overall, these tasks were clearly difficult for children in comparison to similar previous studies. However, it is not clear whether this difficulty was most attributable to the design of the tasks themselves, the length of laboratory visits including multiple experiments, or possibly even to the broader impact of the COVID-19 pandemic. Children who participated in these studies were born between 2018-2020 in the autistic group and 2019-2021 in the NT group, and as such many of these children spent much of their early postnatal (or even perinatal) period under public health restrictions which limited opportunities for social interaction outside the home and increased stress for many families. Several studies have reported subtle developmental delays associated with the pandemic among some children (Deoni et al., 2021; Imboden et al., 2022; Sato et al., 2023; but see Kartushina et al., 2021). Children also participated in other experimental eye-gaze tasks in their laboratory visits in addition to those reported in this dissertation, possibly increasing fatigue and cognitive load. As discussed previously, these studies also differed in various aspects of experimental design in comparison to other studies that may have increased the difficulty for young children. In particular, test trials were relatively short and included only one repetition of the target word.

Whatever the reason, autistic children tested at the Year 2 time point who had sufficient data for inclusion in analyses did not as a group demonstrate average accuracy above chance in the CSL-NV and O+V tasks. It may therefore be argued that investigating individual differences within the autistic group in these tasks captures random noise rather than meaningful variation. As I noted in Chapter 5, a large portion of children were individually above chance in absolute terms, and the wide variation in performance in these tasks is more representative of the heterogeneity of the autism spectrum because I did not limit the sample based on accuracy or ability. I also believe it is important to include children who did not learn in the task such that the individual difference analyses provide insight into characteristics predictive of the full range of possible learning outcomes. Still, future studies may benefit from ensuring group-level learning prior to proceeding with individual differences analyses.

Another limitation related to experimental design was that the linguistic and visual stimuli were not counterbalanced across tasks; that is, each of the CSL-NV, CSL+V, and O+V tasks included different sets of words and images. This decision was in part due to the within-subjects design. The experimental protocol at the Year 2 time point included all three tasks, so the same stimuli could not be repeated. The fact that there were three tasks also complicated the counterbalancing as there were only two experimental orders administered. Finally, the available stimuli were also limited because most children were administered other experimental tasks using stimuli from the same database during their laboratory visits as part of the larger parent study. Thus, any comparison across tasks could be confounded by differences in stimuli. However, it is worth noting again that modeling did not reveal any significant differences in autistic children's performance when comparing the CSL-NV task and the CSL+V task or the CSL+V task and the O+V task. It is recommended that future studies counterbalance stimuli across conditions.

Another limitation of these studies in terms of design and data analysis was the relatively small number of trials, which after aggregating at the child level resulted in even fewer observations that were ultimately included in analyses. Each experiment had originally included 18 test trials (2 familiar, 8 trained (i.e., fast mapping), and 8 untrained (i.e., generalization). The familiar test trials were included only to familiarize children with the structure of the test phase and maintain attention and were therefore not analyzed, while the experimental design error on the untrained (generalization) trials precluded their inclusion in analyses. Thus, each participant could only contribute a possible total of 8 test trials per task. I opted to average the accuracy data across these trials at the participant level such that there was only one observation per participant per task (if they provided sufficient data for inclusion). I chose this approach in part to align with Venker (2019) to facilitate comparison across studies but also to meet the assumptions of the t-tests and then keep the data as consistent as possible across the other analyses. However, more observations would have been preferable analytically (Brauer & Curtin, 2018). Future studies may consider analyzing trial-level data and accounting for the resulting non-independence using mixed-models rather than aggregating. Another way similar data might be analyzed differently in future studies is examining target looking in relation to time. While I did not have any specific hypotheses related to the time course of looking in these studies, Hartley et al. (2020) reported that autistic children were slower in identifying named referents than NT peers in a CSL task, so examining reaction time or using a growth curve modeling approach could yield interesting insights in future work.

Finally, it is critical to acknowledge that the operationalization of learning in these studies is very limited, reflecting only participants' immediate selection of referents to which they had previous exposure during the training phase, often referred to as encoding in the memory literature (e.g., Vlach, 2019). However, learning a novel word involves not only encoding a mapping, but also retaining and retrieving that mapping after delays and interference and generalizing to new lexical category members. Indeed, according to some researchers “learning” per se can only be measured after a delay (e.g., Bjork & Bjork, 2020). Measuring learning only at the encoding phase may obscure more lasting learning and meaningful effects as conditions that support performance during training may not necessarily be the same conditions that facilitate retention, retrieval, or generalization (Schmidt & Bjork, 1992). In the present studies, I originally intended to measure generalization of learning during the test phase by presenting participants with a novel named exemplar that matched training exemplars in shape but differed in color. However, the distractor image in each generalization test trial was not a novel color, confounding children’s looking behavior based on learning with looking based on novelty. As such, those trials were not interpretable. One possible additional consequence of this experimental design error I considered was that participants might have picked up on the pattern of target images having a novel color and used that regularity to inform their looking even on fast mapping trials. To examine this possibility, I visualized children’s accuracy on the first half of test trials compared to the second half of test trials. Similar to the main analyses, to increase the reliability of the average accuracy measures, I included participants only if they contributed >25% of total possible trials for a given half of the test phase (i.e., more than 1 trial out of a possible total of 4). Based on visual inspection, the resulting plots for the CSL-NV task (Figure 14), CSL+V task (Figure 15), and O+V task (Figure 16) do not appear to follow any clear overall pattern suggesting children picked up on this unintentional cue. Because this experimental design error unfortunately precluded the present studies from speaking to generalization, future research will be needed assess generalization of learned words to novel exemplars as well as retention and retrieval of learned words after a delay.

Overall Conclusions

The broad aim of this dissertation was to better understand how young autistic children learn words across various conditions that characterize naturalistic input. Specifically, I investigated children’s ability to learn words amidst referential ambiguity (i.e., CSL) and perceptual variability. Results painted a complex and somewhat surprising picture; I found that autistic children did not show evidence of CSL in a basic task pairing one exemplar image with each novel word throughout training but – contrary to expectations – did demonstrate learning in a CSL task which paired each novel word with three exemplars that shared the same shape but varied in color during training, and outperformed younger NT children on this task when covarying nonverbal cognitive ability. Also contrary to expectations, autistic children’s learning of novel words referring to multiple perceptually variable exemplars was not improved by more explicit naming, and the extent to which explicit naming improved performance in the tasks including variable exemplars was larger in NT children compared to autistic children when covarying nonverbal cognitive ability. Results may point to the influence of perceptually salient stimuli in autistic children’s word learning, though more research will be needed to examine this possibility. Finally, investigation of individual differences predictive of autistic children’s performance in these tasks revealed a link between CSL and receptive language when accounting for nonverbal cognitive ability, level of autism traits, and chronological age, and a link between chorological age and performance in the task that included explicitly named variable exemplars when accounting for nonverbal cognitive ability, receptive language, and level of autism traits. While cautious interpretation is recommended considering the outlined study limitations, the findings of this dissertation provide preliminary insights into mechanisms underlying word learning in young autistic children across multiple conditions. In doing so, this work extends the reach of statistical learning theory and may inform future language interventions and environmental accommodations tailored more specifically for autistic children based on their individual profile of strengths and needs.

References

Ahufinger, N., Guerra, E., Ferinu, L., Andreu, L., & Sanz-Torrent, M. (2021). Cross-situational statistical learning in children with developmental language disorder. *Language, Cognition and Neuroscience*, *36*(9), 1180–1200. https://doi.org/10.1080/23273798.2021.1922723

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347–358. https://doi.org/10.1177/014272379901905703

Alt, M., Meyers, C., Oglivie, T., Nicholas, K., & Arizmendi, G. (2014). Cross-situational statistically based word learning intervention for late-talking toddlers. *Journal of Communication Disorders*, *52*, 207–220. https://doi.org/10.1016/j.jcomdis.2014.07.002

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, *9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Broedelet, I., Boersma, P., & Rispens, J. (2023). Implicit cross-situational word learning in children with and without developmental language disorder and its relation to lexical-semantic knowledge. *Frontiers in Communication*, *8*. https://doi.org/10.3389/fcomm.2023.1021654

Chien, Y.-L., Gau, S. S.-F., Chiu, Y.-N., Tsai, W.-C., Shang, C.-Y., & Wu, Y.-Y. (2014). Impaired sustained attention, focused attention, and vigilance in youths with autistic disorder and Asperger's disorder. *Research in Autism Spectrum Disorders*, *8*(7), 881–889. https://doi.org/10.1016/j.rasd.2014.04.006

Chien, Y.-L., Gau, S. S.-F., Shang, C.-Y., Chiu, Y.-N., Tsai, W.-C., & Wu, Y.-Y. (2015). Visual memory and sustained attention impairment in youths with autism spectrum disorders. *Psychological Medicine*, *45*(11), 2263–2273. https://doi.org/10.1017/S0033291714003201

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*(3), 331–343. https://doi.org/10.1017/S1355617709090481

Deoni, S. C., Beauchemin, J., Volpe, A., D'Sa, V., & Consortium, the R. (2021). *Impact of the COVID-19 Pandemic on Early Child Cognitive Development: Initial Findings in a Longitudinal Observational Study of Child Health* (p. 2021.08.10.21261846). medRxiv. https://doi.org/10.1101/2021.08.10.21261846

Girolamo, T., Shen, L., Monroe-Gulick, A., Rice, M. L., & Eigsti, I.-M. (2023). Studies pertaining to language impairment in school-age autistic individuals underreport participant socio-demographics: A systematic review. *Autism*, 13623613231166749. https://doi.org/10.1177/13623613231166749

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*(July 2019), 104265. https://doi.org/10.1016/j.cognition.2020.104265

Howlin, P. (2003). Outcome in High-Functioning Adults with Autism with and Without Early Language Delays: Implications for the Differentiation Between Autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, *33*(1).

Imboden, A., Sobczak, B. K., & Griffin, V. (2022). The impact of the COVID-19 pandemic on infant and toddler development. *Journal of the American Association of Nurse Practitioners*, *34*(3), 509–519. https://doi.org/10.1097/JXX.0000000000000653

Johnson, K. A., Robertson, I. H., Kelly, S. P., Silk, T. J., Barry, E., Dáibhis, A., Watchorn, A., Keavey, M., Fitzgerald, M., Gallagher, L., Gill, M., & Bellgrove, M. A. (2007). Dissociation in performance of children with ADHD and high-functioning autism on a task of sustained attention. *Neuropsychologia*, *45*(10), 2234–2245. https://doi.org/10.1016/j.neuropsychologia.2007.02.019

Kartushina, N., Mani, N., Aktan-Erciyes, A., Alaslani, K., Aldrich, N. J., Almohammadi, A., Alroqi, H., Anderson, L., Andonova, E., Aussems, S., Babineau, M., Barokova, M., Bergmann, C., Cashon, C. H., Custode, S., De Carvalho, A., Dimitrova, N., Dynak, A., Farah, R., … Mayor, J. (2021). *COVID-19 first lockdown as a window into language acquisition: Associations between caregiver-child activities and vocabulary gains.* https://doi.org/10.31234/osf.io/5ejwu

Little, L. M., Dean, E., Tomchek, S. D., & Dunn, W. (2017). Classifying sensory profiles of children in the general population. *Child: Care, Health and Development*, *43*(1), 81–88. https://doi.org/10.1111/cch.12391

Mayo, J., Chlebowski, C., Fein, D. A., & Eigsti, I.-M. (2013). Age of First Words Predicts Cognitive Ability and Adaptive Skills in Children with ASD. *Journal of Autism and Developmental Disorders*, *43*(2), 253–264. https://doi.org/10.1007/s10803-012-1558-0

McGregor, K. K., Rost, G., Arenas, R., Farris-Trimble, A., & Stiles, D. (2013). Children with ASD can use gaze in support of word recognition and learning. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(7), 745–753. https://doi.org/10.1111/jcpp.12073

McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbisi-Kelm, T., & Pomper, R. (2022). What Children with Developmental Language Disorder Teach Us About Cross-Situational Word Learning. *Cognitive Science*, *46*(2). https://doi.org/10.1111/cogs.13094

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48. https://doi.org/10.1037/0021-843X.110.1.40

Pugh, S. L., Brown, P. D., & Enserro, D. (2021). Missing repeated measures data in clinical trials. *Neuro-Oncology Practice*, *9*(1), 35–42. https://doi.org/10.1093/nop/npab043

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sato, K., Fukai, T., Fujisawa, K. K., & Nakamuro, M. (2023). Association Between the COVID-19 Pandemic and Early Childhood Development. *JAMA Pediatrics*, *177*(9), 930. https://doi.org/10.1001/jamapediatrics.2023.2096

Schmidt, R. A., & Bjork, R. A. (1992). New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychological Science*, *3*(4), 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x

Smith, L. B. (1999). Children’s noun learning: How general learning processes make specialized learning mechanisms. In *The emergence of language* (pp. 277–303).

Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to cimpian and markman (2005) and booth, Waxman, and Huang (2005). *Developmental Psychology*, *42*(6), 1339–1343. https://doi.org/10.1037/0012-1649.42.6.1339

Smith, L. B., & Yu, C. (2013). Visual Attention Is Not Enough: Individual Differences in Statistical Word-Referent Learning in Infants. *Language Learning and Development*, *9*(1), 25–49. https://doi.org/10.1080/15475441.2012.707104

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

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

Tovar, Á. E., Rodríguez-Granados, A., & Arias-Trejo, N. (2020). Atypical shape bias and categorization in autism: Evidence from children and computational simulations. *Developmental Science*, *23*(2), 1–14. https://doi.org/10.1111/desc.12885

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*(November 2018), 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Venker, C. E., Mathée, J., Neumann, D., Edwards, J., Saffran, J., & Ellis Weismer, S. (2021). Competing Perceptual Salience in a Visual Word Recognition Task Differentially Affects Children With and Without Autism Spectrum Disorder. *Autism Research*, *14*(6), 1147–1162. https://doi.org/10.1002/aur.2457

Venker, C. E., Neumann, D., & Aladé, F. (2022). Visual perceptual salience and novel referent selection in children with and without autism spectrum disorder. *Autism &*

Developmental Language Impairments, *7*, 239694152210854. https://doi.org/10.1177/23969415221085476

Vivanti, G., Fanning, P. A. J., Hocking, D. R., Sievers, S., & Dissanayake, C. (2017). Social Attention, Joint Attention and Sustained Attention in Autism Spectrum Disorder and Williams Syndrome: Convergences and Divergences. *Journal of Autism and Developmental Disorders*, *47*(6), 1866–1877. https://doi.org/10.1007/s10803-017-3106-4

Vlach, H. A. (2019). Learning to Remember Words: Memory Constraints as Double-Edged Sword Mechanisms of Language Development. *Child Development Perspectives*, *13*(3), 159–165. https://doi.org/10.1111/cdep.12337

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children’s cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230. https://doi.org/10.1016/j.jml.2016.10.001

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children’s retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 700–711. https://doi.org/10.1037/xlm0000611.Statistics

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants’ cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015.Memory

Yerys, B. E., Wallace, G. L., Sokoloff, J. L., Shook, D. A., James, J. D., & Kenworthy, L. (2009). Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. *Autism Research: Official Journal* *of the International Society for Autism Research*, *2*(6), 322–333. https://doi.org/10.1002/aur.103

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales Fifth Edition (PLS-5)*. Pearson.

Tables and Figures

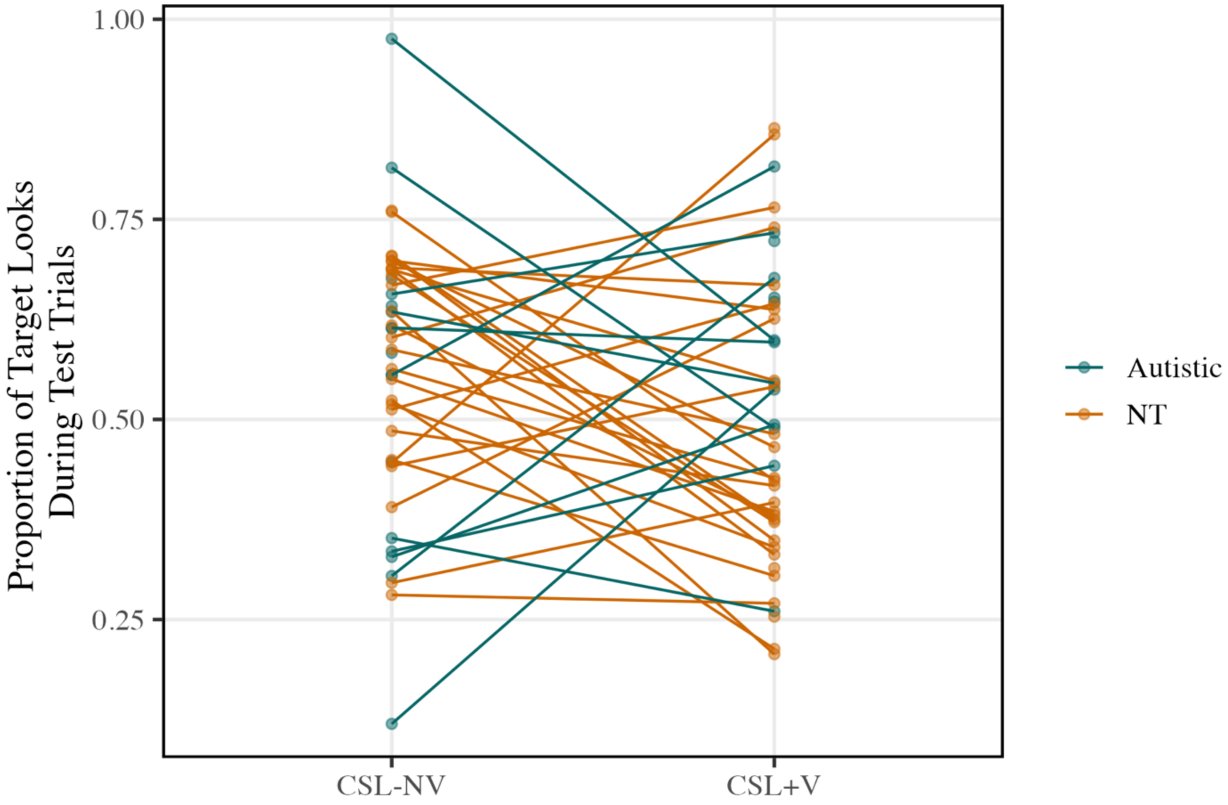


Figure 14. Spaghetti plot connecting the average accuracies for individual participants with sufficient data for inclusion in the CSL-NV and CSL+V task. Dots without connecting lines represent the average accuracy for autistic (green) and NT (orange) participants who contributed sufficient data for inclusion in one task but not the other.

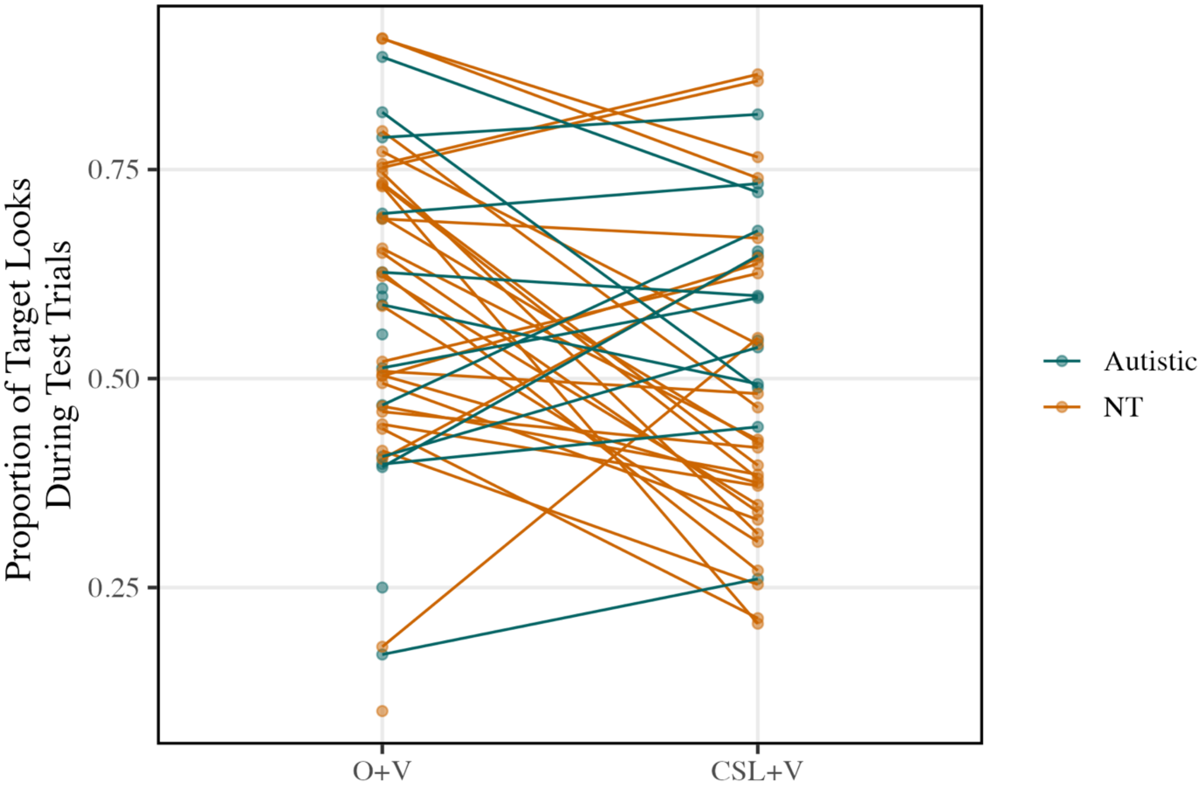


Figure 15. Spaghetti plot connecting the average accuracies for individual participants with sufficient data for inclusion in the O+V and CSL+V task. Dots without connecting lines represent the average accuracy for autistic (green) and NT (orange) participants who contributed sufficient data for inclusion in one task but not the other.

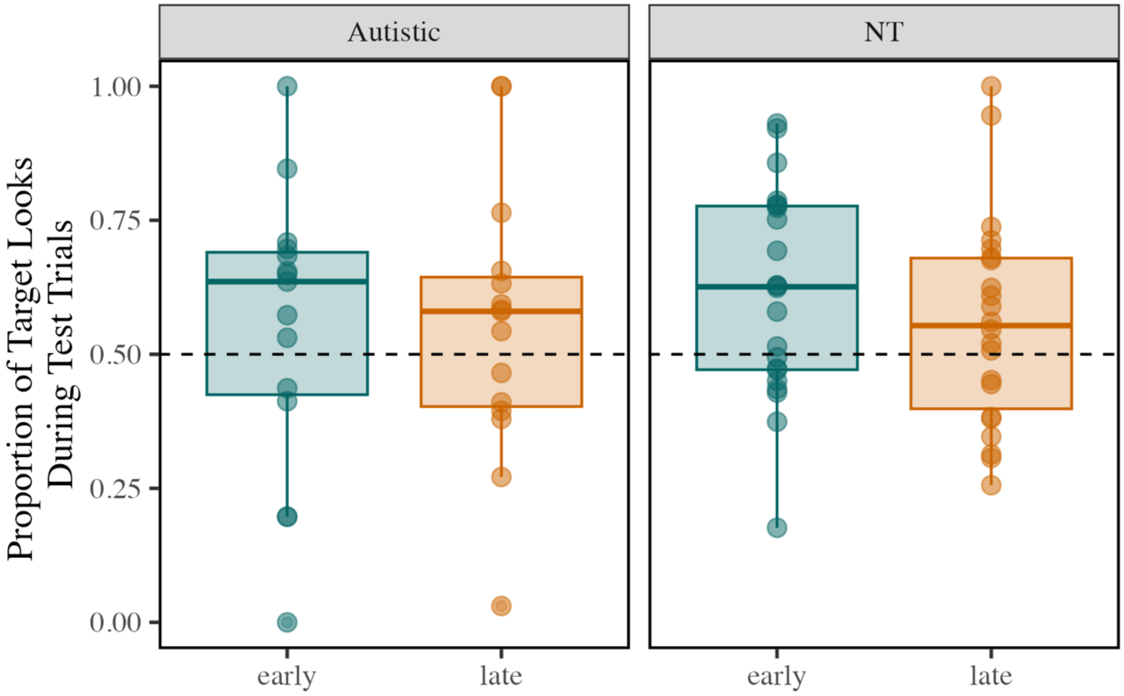


Figure 16. Average accuracy in the first half (“early”; green) and second half (“late”; orange) of fast-mapping test trials for participants with sufficient data for inclusion (in the age range of the Year 2 time point of the larger study) in the CSL-NV task for autistic children (left) and NT children (right).

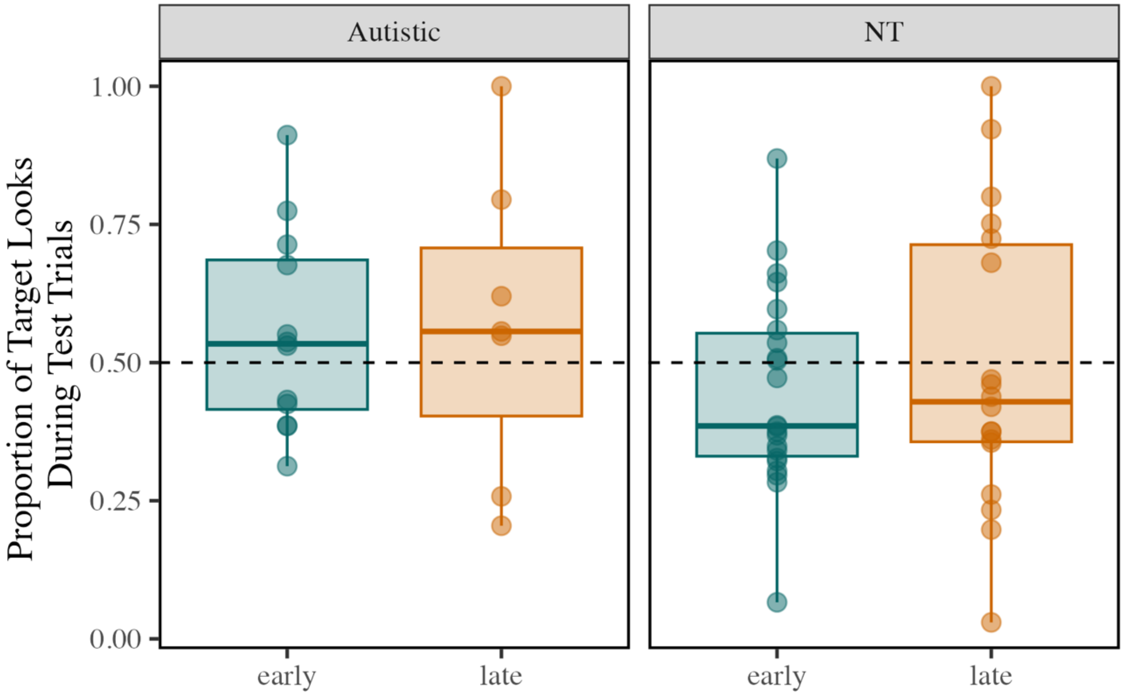


Figure 17. Average accuracy in the first half (“early”; green) and second half (“late”; orange) of fast-mapping test trials for participants with sufficient data for inclusion in the CSL+V task for autistic children (left) and NT children (right).

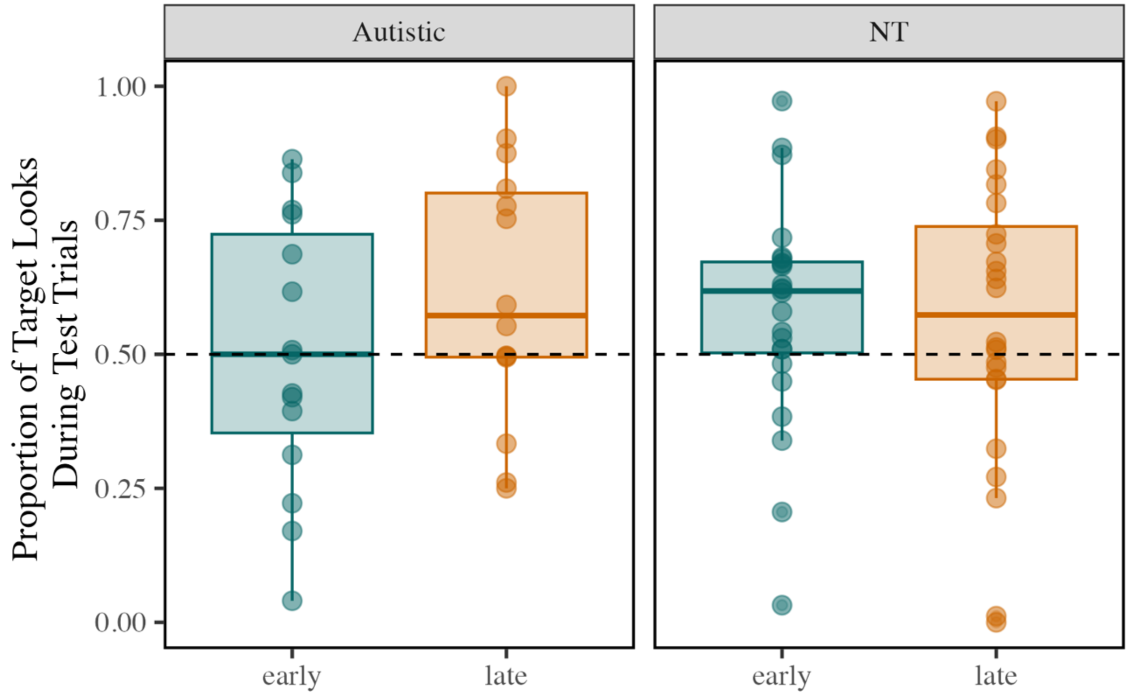


Figure 18. Average accuracy in the first half (“early”; green) and second half (“late”; orange) of fast-mapping test trials for participants with sufficient data for inclusion in the O+V task for autistic children (left) and NT children (right).

Appendix A

Novel visual stimuli appearing during training in experimental eye-gaze tasks. Note that each object or set of objects were given a different novel label in each counterbalanced version (Order A or Order B) of the task. Taken from the *NOUN* Database. Some alterations made in Adobe Photoshop.

CSL-NV Task:

Hux (Order A) or *Zeb* (Order B)

Tife (Order A) or *Jick* (Order B)

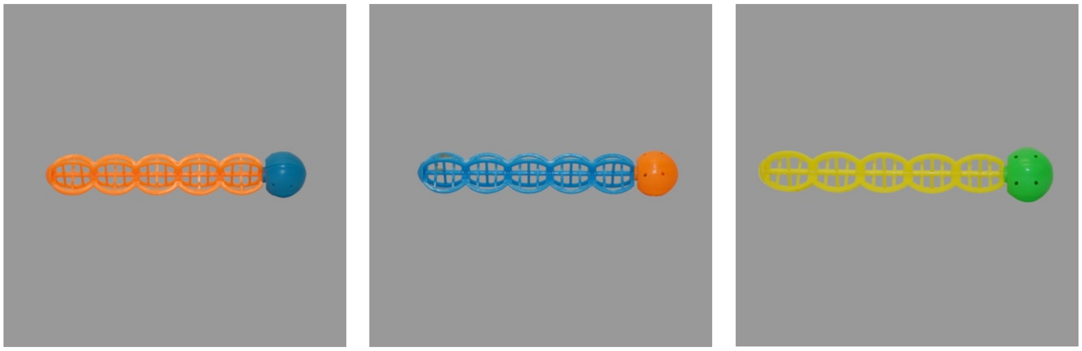
Jick (Order A) or *Tife* (Order B)

Zeb (Order A) or *Hux* (Order B)

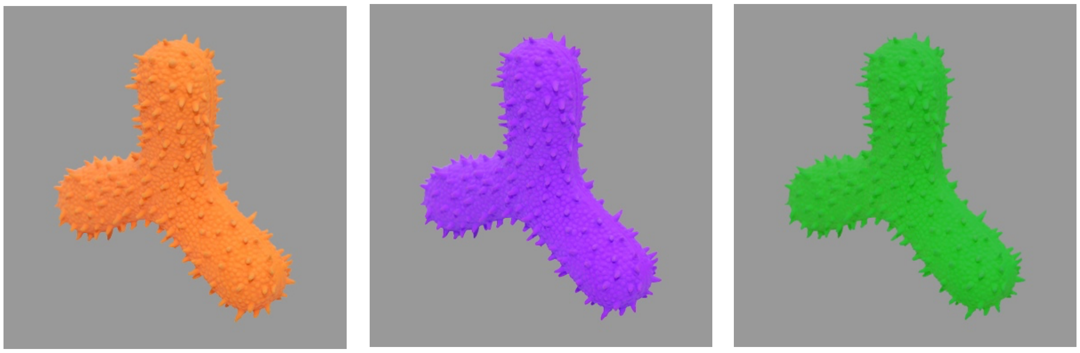
CSL+V Task:

Dax (Order A) or *Gip* (Order B)

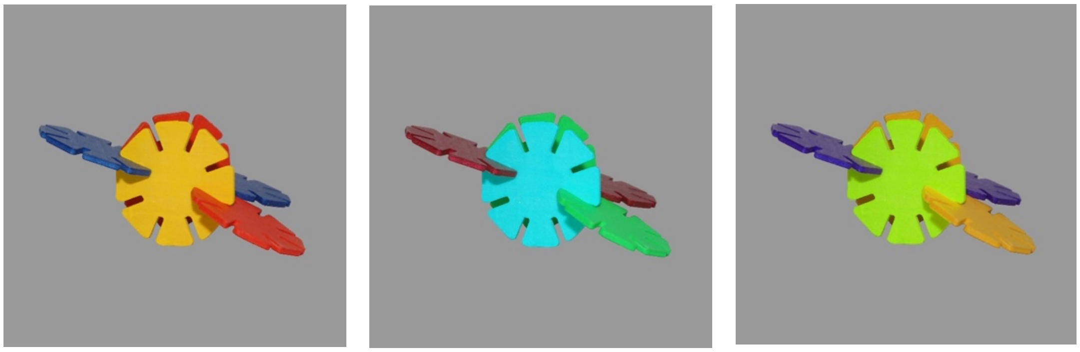
Gip (Order A) or *Dax* (Order B)



Mel (Order A) or *Poss* (Order B)



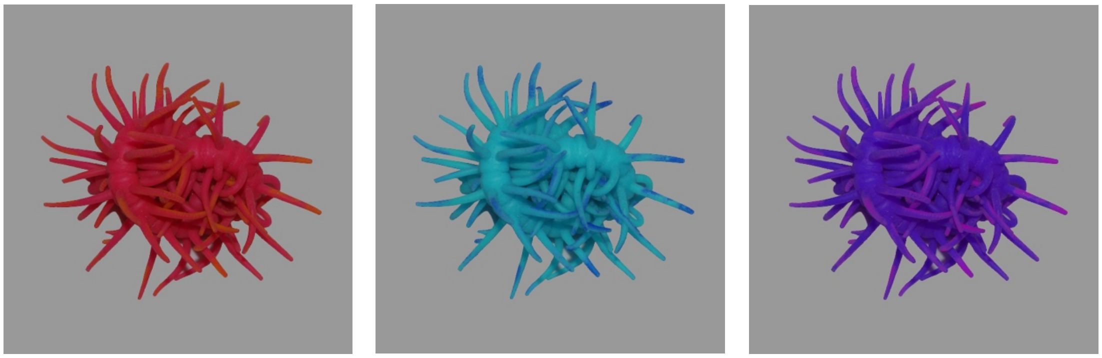
Poss (Order A) or *Mel* (Order B)



O+V Task:

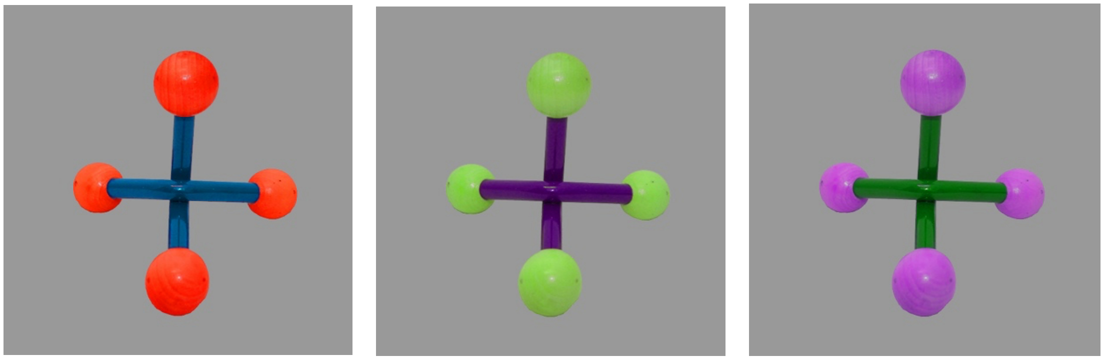
Dage (Order A) or *Shede* (Order B)

Juff (Order A) or *Lep* (Order B)



Lep (Order A) or *Juff* (Order B)

Shede (Order A) or *Dage* (Order B)



Appendix B

Counterbalanced trial orders (Order A and Order B) for each experimental eye-gaze task based on Venker (2019) with some modifications. Words listed are the novel **labels** given to the referents. Images paired with those labels were counterbalanced between Order A and Order B (see Appendix A). In test trials, target images appeared side-by-side with a distractor image which had also appeared during training and varied across trials (i.e., target and distractor images were not yoked).

CSL-NV Task Order A:

Trial	Trial Phase	Word 1	Word 2	Sequential	Target Word	Target Side
1	pre-train	car	book	No		
2	pre-train	dog	book	Yes		
3	pre-train	car	dog	Yes		
4	pre-train	book	dog	No		
5	train	zeb	hux	No		
6	train	tife	jick	Yes		
7	train	jick	hux	Yes		
8	train	hux	zeb	No		
9	train	tife	zeb	Yes		
10	train	zeb	jick	No		
11	train	jick	hux	No		
12	train	hux	tife	Yes		
13	train	zeb	tife	Yes		
14	train	jick	zeb	No		
15	train	jick	zeb	Yes		
16	train	hux	tife	No		
17	train	hux	jick	Yes		
18	train	tife	jick	No		
19	train	tife	hux	No		
20	train	hux	zeb	Yes		
21	train	zeb	tife	No		
22	train	tife	jick	Yes		
23	train	zeb	hux	Yes		
24	train	jick	tife	No		
25	test_familiar				car	L
26	test_trained				jick	R
27	test_untrained				tife	L
28	test_untrained				zeb	R
29	test_trained				hux	R
30	test_trained				tife	L
31	test_untrained				jick	R

32	test_untrained				hux	R
33	test_trained				zeb	L
34	test_familiar				book	L
35	test_trained				jick	R
36	test_untrained				tife	R
37	test_untrained				zeb	L
38	test_trained				hux	L
39	test_trained				zeb	R
40	test_trained				tife	L
41	test_untrained				hux	R
42	test_untrained				jick	L

CSL-NV Task Order B:

Trial	Trial Phase	Word 1	Word 2	Sequential	Target Word	Target Side
1	pre-train	car	dog	Yes		
2	pre-train	car	book	No		
3	pre-train	book	dog	No		
4	pre-train	dog	book	Yes		
5	train	tife	hux	Yes		
6	train	zeb	jick	No		
7	train	zeb	tife	Yes		
8	train	jick	tife	No		
9	train	jick	zeb	No		
10	train	hux	zeb	No		
11	train	jick	tife	Yes		
12	train	tife	zeb	Yes		
13	train	zeb	hux	No		
14	train	jick	hux	Yes		
15	train	zeb	hux	Yes		
16	train	hux	jick	No		
17	train	jick	tife	Yes		
18	train	hux	zeb	Yes		
19	train	tife	jick	No		
20	train	hux	tife	No		
21	train	tife	zeb	No		
22	train	zeb	jick	Yes		
23	train	hux	jick	Yes		
24	train	tife	hux	No		
25	test_familiar				dog	L
26	test_trained				zeb	R
27	test_trained				tife	L

28	test_untrained				hux	R
29	test_untrained				jick	L
30	test_untrained				zeb	R
31	test_trained				hux	L
32	test_untrained				tife	R
33	test_trained				jick	R
34	test_familiar				dog	L
35	test_trained				hux	R
36	test_trained				zeb	L
37	test_untrained				jick	R
38	test_untrained				hux	R
39	test_untrained				tife	L
40	test_untrained				zeb	L
41	test_trained				tife	R
42	test_trained				jick	L

CSL+V Task Order A:

Trial	Trial Type	Word 1	Word 2	Sequential	Target Word	Target Side
1	pre-train	ball	cup	Yes		
2	pre-train	ball	shoe	No		
3	pre-train	shoe	cup	No		
4	pre-train	cup	shoe	Yes		
5	train	poss	dax	Yes		
6	train	mel	gip	No		
7	train	mel	poss	Yes		
8	train	gip	poss	No		
9	train	gip	mel	No		
10	train	dax	mel	No		
11	train	gip	poss	Yes		
12	train	poss	mel	Yes		
13	train	mel	dax	No		
14	train	gip	dax	Yes		
15	train	mel	dax	Yes		
16	train	dax	gip	No		
17	train	gip	poss	Yes		
18	train	dax	mel	Yes		
19	train	poss	gip	No		
20	train	dax	poss	No		

21	train	poss	mel	No		
22	train	mel	gip	Yes		
23	train	dax	gip	Yes		
24	train	poss	dax	No		
25	test_familiar				cup	L
26	test_trained				mel	R
27	test_trained				poss	L
28	test_untrained				dax	R
29	test_untrained				gip	L
30	test_untrained				mel	R
31	test_trained				dax	L
32	test_untrained				poss	R
33	test_trained				gip	R
34	test_familiar				cup	L
35	test_trained				dax	R
36	test_trained				mel	L
37	test_untrained				gip	R
38	test_untrained				dax	R
39	test_untrained				poss	L
40	test_trained				gip	L
41	test_trained				poss	R
42	test_untrained				mel	L

CSL+V Task Order B:

Trial	Trial Type	Word 1	Word 2	Sequential	Target Word	Target Side
1	pre-train	ball	shoe	No		
2	pre-train	cup	shoe	Yes		
3	pre-train	ball	cup	Yes		
4	pre-train	shoe	cup	No		
5	train	poss	mel	No		
6	train	gip	dax	Yes		
7	train	dax	mel	Yes		
8	train	mel	poss	No		
9	train	gip	poss	Yes		
10	train	poss	dax	No		
11	train	dax	mel	No		
12	train	mel	gip	Yes		

13	train	poss	gip	Yes		
14	train	dax	poss	No		
15	train	dax	poss	Yes		
16	train	mel	gip	No		
17	train	mel	dax	Yes		
18	train	gip	dax	No		
19	train	gip	mel	No		
20	train	mel	poss	Yes		
21	train	poss	gip	No		
22	train	gip	dax	Yes		
23	train	poss	mel	Yes		
24	train	dax	gip	No		
25	test_familiar				ball	L
26	test_trained				dax	L
27	test_untrained				gip	L
28	test_untrained				poss	R
29	test_trained				mel	R
30	test_trained				gip	L
31	test_untrained				dax	R
32	test_untrained				mel	R
33	test_trained				poss	L
34	test_familiar				cup	L
35	test_trained				dax	R
36	test_untrained				gip	R
37	test_untrained				poss	L
38	test_trained				mel	L
39	test_trained				poss	R
40	test_trained				gip	R
41	test_untrained				mel	L
42	test_untrained				dax	R

O+V Task Order A:

Trial	Trial Type	Target Word	Target Side
1	pre-train	cup	
2	pre-train	shoe	
3	pre-train	ball	
4	pre-train	cup	

5	train	dage	
6	train	juff	
7	train	dage	
8	train	juff	
9	train	lep	
10	train	juff	
11	train	dage	
12	train	shede	
13	train	lep	
14	train	shede	
15	train	juff	
16	train	shede	
17	train	dage	
18	train	lep	
19	train	dage	
20	train	juff	
21	train	shede	
22	train	lep	
23	train	dage	
24	train	shede	
25	train	lep	
26	train	dage	
27	train	juff	
28	train	shede	
29	train	juff	
30	train	lep	
31	train	shede	
32	train	juff	
33	train	dage	
34	train	lep	
35	train	juff	
36	train	dage	
37	train	lep	
38	train	shede	
39	train	dage	
40	train	lep	
41	train	shede	
42	train	lep	
43	train	shede	

44	train	juff	
45	test_familiar	cup	L
46	test_trained	juff	R
47	test_untrained	dage	L
48	test_untrained	lep	R
49	test_trained	shede	L
50	test_untrained	juff	R
51	test_trained	lep	L
52	test_trained	dage	R
53	test_untrained	shede	R
54	test_familiar	cup	L
55	test_untrained	lep	R
56	test_trained	juff	L
57	test_trained	shede	R
58	test_untrained	dage	L
59	test_trained	lep	R
60	test_trained	dage	R
61	test_untrained	juff	L
62	test_untrained	shede	L

O+V Task Order B:

Trial	Trial Type	Target Word	Target Side
1	pre-train	ball	
2	pre-train	cup	
3	pre-train	shoe	
4	pre-train	ball	
5	train	lep	
6	train	juff	
7	train	shede	
8	train	juff	
9	train	shede	
10	train	dage	
11	train	shede	
12	train	lep	
13	train	juff	
14	train	shede	
15	train	lep	

16	train	juff	
17	train	shede	
18	train	dage	
19	train	shede	
20	train	lep	
21	train	dage	
22	train	juff	
23	train	dage	
24	train	shede	
25	train	dage	
26	train	juff	
27	train	dage	
28	train	shede	
29	train	lep	
30	train	juff	
31	train	lep	
32	train	dage	
33	train	lep	
34	train	dage	
35	train	juff	
36	train	dage	
37	train	juff	
38	train	lep	
39	train	juff	
40	train	lep	
41	train	shede	
42	train	dage	
43	train	lep	
44	train	shede	
45	test_familiar	ball	L
46	test_trained	juff	R
47	test_untrained	shede	L
48	test_untrained	dage	R
49	test_trained	lep	R
50	test_trained	shede	L
51	test_untrained	juff	R
52	test_untrained	lep	R
53	test_trained	dage	L
54	test_familiar	shoe	L

55	test_untrained	shede	R
56	test_untrained	dage	L
57	test_trained	juff	R
58	test_trained	lep	L
59	test_trained	dage	R
60	test_trained	shede	R
61	test_untrained	lep	L
62	test_untrained	juff	L